Implantação e gestão de repositórios institucionais:

políticas, memória, livre acesso e preservação

Implantação e gestão de repositórios institucionais:

políticas, memória, livre acesso e preservação

ORGANIZADORES

Luis Sayão

Lídia Brandão Toutain

Flavia Garcia Rosa

Carlos Henrique Marcondes

EDUFBA

Salvador

2009

Sumário

Apresentação

A segunda metade do século XX foi um período marcante em transformações decorrentes dos sistemas de comunicação e informação. Estas alterações trouxeram como resultados a aplicabilidade desses sistemas na economia, na política e na cultura. Há um deslocamento do regime de acumulação baseado na grande indústria, para as atividades imateriais, deslocamento este, imposto pela globalização do mercado com relação à produção intelectual e alterando, também as relações sociais.

Os avanços tecnológicos da comunicação e da informação (TIC) permitiram o surgimento de redes de comunicação eletrônica, revolucionando os fluxos de informação, forma de acesso e troca de informações ampliando o espiral do conhecimento, graças a um novo parâmetro espaço-tempo possibilitado pelas tecnologias. A transmissão de dados ocorre instantaneamente graças à convergência tecnológica com as telecomunicações e com a micro-eletrônica. Esses são os instrumentos e as relações que servem de base para a Sociedade da Informação (SI).

O século XXI consolida esse formato da informação – eletrônica. A internet possibilitou acesso às fontes de informação sendo relevante o correio eletrônico, a *Web,* listas de discussão – permitem a troca de informações entre os especialistas, *blogs*, *chats* – a disponibilização de versões preliminares de artigos de periódicos e de comunicações em conferências científicas, de capítulos de livros, publicações periódicas electrônicas materiais de ensino bases de dados, compilações de dados electrônicos, e os repositórios, dentre outros. Essas mudanças influenciarão, sobremaneira, a disseminação da produção científica e as relações de comunicação da informação no ambiente acadêmico.

A ênfase nesse livro editado pela Editora da Universidade Federal da Bahia são os Repositórios Institucionais cuja finalidade é gerenciar e ampliar a visibilidade da informação científica, artística e cultural produzida pelas instituições. Essa visibilidade que promove uma salutar competitividade entre elas permite o acesso e a divulgação,

com maior rapidez, às conquistas e inovações nas áreas artísticas e técnico-científicas, através da gestão do patrimônio documental e iconográfico relativo à produção dessas instituições.

A estrutura do livro contempla questões que vão dos softwares livres a preservação digital, passando por abordagens sobre a preservação da memória, arquitetura da informação, recuperação da informação, insumos, preservação digital. O texto introdutório assinado por Carlos Henrique Marcondes e Luis Fernando Sayão contextualiza a temática central do livro – repositórios institucionais e acesso livre – de forma ampla e atual dentro dos novos paradigmas da comunicação científica. Finalmente, a última parte do livro é enriquecida com relatos de experiências que com certeza contribuirão para que outras instituições de ensino superior trilhem esse caminho da disponibilização de conteúdos de forma ampla e sem barreiras, pautada na Iniciativa de Arquivos Abertos e no Movimento de Acesso Livre.

Flávia Garcia Rosa
Lidia Brandão Toutain

À guisa de introdução:
repositórios institucionais e livre acesso

Carlos Henrique Marcondes
Luis Fernando Sayão

> Uma versão completa da obra e todos os materiais suplementares, incluindo uma cópia da licença, como acima definida, é depositada e, portanto, publicada em um formato eletrônico normalizado e apropriado em pelo menos um repositório que utilize normas técnicas adequadas (como as definições estabelecidas pelo modelo Open Archives) e que seja mantido por uma instituição acadêmica, sociedade científica, organismo governamental, ou outra organização estabelecida que pretenda promover o acesso livre, a distribuição irrestrita, a interoperabilidade e o arquivamento a longo prazo. (IBICT, 2005).

O texto anterior, retirado do *Manifesto Brasileiro de Apoio ao Acesso Livre à Informação Científica,* lançado pelo Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) em 13 de setembro de 2005, consagra o papel dos repositórios institucionais, objeto deste livro, como peças fundamentais de um futuro sistema brasileiro de livre acesso à produção científica do país. Um repositório institucional é uma biblioteca digital destinada a guardar, preservar e garantir livre acesso, via internet, à produção científica no âmbito de uma dada instituição.

No entanto, muito mais que uma peça tecnológica, um repositório institucional se constitui hoje, no contexto de um amplo e crescente movimento internacional de apoio ao livre acesso à informação científica, num ator político (RIEGER, 2008) que desempenha um papel inédito, como nunca antes visto, no ciclo de comunicação científica.

O mundo emergiu da Segunda Guerra Mundial com uma crescente

percepção do papel central da Ciência para o desenvolvimento e, dentro disso, do papel estratégico da Informação em Ciência e Tecnologia (ICT) como insumo para a ciência. Gerir e otimizar os fluxos da ICT se tornam, ao longo das décadas de 60 e 70 do século XX, atividades estratégicas de governos e organismo internacionais. Em uma das primeiras tentativas de identificar e mapear diferentes fluxos informacionais, os diversos atores – pessoais, institucionais, políticos, econômicos, tecnológicos etc. –, e seus papéis no nascente ciclo de comunicação científica na década de 1970, o assim chamado modelo UNISIST (1971), não previa a figura dos repositórios institucionais.

Repositórios institucionais são entendidos hoje como elementos de uma rede ou infraestrutura informacional de um país ou de um domínio institucional destinados a garantir a guarda, preservação a longo prazo e, fundamentalmente, o livre acesso à produção científica de uma dada instituição. A lógica que preside o surgimento dos repositórios institucionais no cenário internacional da ICT é a retomada de uma proposta que tem suas raízes no Iluminismo: os resultados da atividade científica, na forma das diferentes publicações, resultados estes muitas vezes obtidos à custa de pesados investimentos públicos, devem necessariamente também ser públicos, poder ser utilizados amplamente, não serem apropriados de forma privada. Assim, cada instituição científica ou acadêmica, e sua correspondente comunidade, deve manter em seu repositório institucional de livre acesso cópias da produção científica de sua comunidade.

Onde então, quando, como e em que contexto surgem no cenário dos fluxos de ICT os repositórios institucionais?

O lançamento do *Manifesto brasileiro de apoio ao acesso livre à informação científica* em 2005 é o primeiro desdobramento político no Brasil de um movimento internacional amplo de apoio ao livre acesso à informação científica. Este movimento tem raízes na situação mundial que evolui a partir da crescente valorização da ICT, já mencionada, que acontece após a Segunda Guerra Mundial. No bojo dos esforços empreendidos por diferentes países – inclusive o Brasil, com a criação do Instituto Brasileiro de Bibliografia e Documentação (IBBD), atual IBICT, em 1954 – para gerir e otimizar os fluxos estratégicos de ICT, cria-se, a

partir da década de 1960, um setor econômico crescentemente forte, a chamada indústria da informação. Essa indústria é formada por todo o tipo de serviços de intermediação entre produtores e consumidores ou usuários de informação científica; em ambos os casos trata-se fundamentalmente de pesquisadores e acadêmicos. Entre estes serviços de intermediação destacam-se os serviços de indexação e resumo, os serviços de acesso às bases de dados, as bibliotecas especializadas e acadêmicas, os editores científicos (de periódicos ou de monografias), os serviços de acesso à chamada "literatura cinzenta" – relatórios técnicos, teses, trabalhos publicados em congressos –, todos claramente identificados no modelo UNISIST.

Junto a estes serviços desenvolvem-se posteriormente, como decorrência das pesquisas sobre tratamento quantitativo da crescente produção bibliográfica mundial, serviços que produziam indicadores estatísticos e permitiam avaliar esta produção, o mais importante dos quais é o Institute for Scientific Information (ISI) –, fundado por Eugene Garfield. Apesar dos questionamentos, o ISI adquire uma proeminência e autoridade crescentes na avaliação dos periódicos científicos internacionais, de tal forma que hoje, quer se goste ou não, quer se critique, é uma referência reconhecida pelos mais diferentes atores envolvidos na gestão da ICT por todo o mundo, exercendo por isso um papel monopolista e único – portanto, altamente perigoso[1] – no cenário da avaliação da ICT.

Um dos segmentos economicamente mais fortes da indústria da informação é constituído pelos editores científicos. Editores científicos editam também livros, mas, no que interessa para esta discussão, editam periódicos. Ainda hoje e nos primórdios da editoração científica, muitos deles são tradicionais editores de um único periódico científico, como é o caso da Royal Society britânica, que edita até hoje o histórico *Philosophical Transactions of the Royal Society*, considerado o primeiro periódico e onde escreveram Newton, Faraday, Boyle e outros. Sociedades científicas como a Royal Society, foram na verdade, os primeiros editores científicos. No entanto, quando da formulação

1 Há algum tempo o ISI é propriedade do grupo editorial privado Thomson-Reuters.

do modelo UNISIST, já existia consolidado o segmento de editores científicos privados, que talvez devessem ser mais apropriadamente chamados de publicadores científicos comerciais.

Um publicador científico reúne e publica periódicos científicos editados por diferentes editores. Além de editá-los, um publicador científico organiza sua distribuição e vende assinaturas. À medida que a ICT ganhava importância estratégica entre os países criava-se um mercado crescente de publicações científicas, que alimentava com novas e atualizadas publicações, pesquisadores de todo o mundo, ansiosos por terem acesso aos mais recentes resultados da pesquisa em suas áreas, o que por sua vez alimentaria suas próprias pesquisas e consequentes publicações acadêmicas. Complementa este sistema mecanismos de avaliação da produção científica mundial, dados pelos estudos bibliométricos realizados pelo ISI e atualizados anualmente.

Completa-se assim um quadro que inclui um conjunto de interesses sinergéticos envolvendo a importância estratégica do desenvolvimento científico para as economias dos diversos países: o interesse e a pressão sobre cientistas e acadêmicos no acesso às publicações atualizadas com os últimos resultados de pesquisa de suas respectivas áreas; os publicadores interessados na venda de assinaturas; de bibliotecas especializadas e acadêmicas, interessadas em otimizar suas aquisições e bem atenderem suas comunidades de usuários; o interesse do ISI em se manter como praticamente a única referência na avaliação da produção científica mundial.

Este quadro começa a se desequilibrar a partir de meados da década de 1970. Frente a uma demanda crescente, o setor econômico dos publicadores, como, aliás, de vários setores da economia capitalista, passa por um processo de concentração econômica e olipolização, com fusões, associações, aquisições de empresas umas pelas outras. De centenas de publicadores científicos existentes nas décadas de 1960 e 1970 têm-se hoje poucos e gigantescos grupos publicadores internacionais, como Elsevier, Emerald, Kruger, Springer etc., e cada um dos quais publica centenas e mesmo milhares de títulos de periódicos científicos, vendem assinaturas em pacotes fechados, compostos de periódicos interessantes, mas outros nem tanto, que o publicador tem interesse em promover.

O resultado disso são os crescentes e cada vez mais proibitivos custos das assinaturas. *O Study on the economic and technical evolution of scientific publications market in Europe* assinala que:

> Nos últimos 30 anos, os preços da subscrição de periódicos científicos têm aumentado regularmente. Entre 1975 e 1995 os aumentos foram de 200 a 300% maior que a inflação. A esse fato foi acompanhado por uma queda nas subscrições contratadas tanto por pesquisadores individuais e quanto pelas bibliotecas cujos orçamentos ficaram pressionados. Na realidade, crescimento dos preços dos periódicos ultrapassaram em muito a evolução dos orçamentos das bibliotecas, que cresceram menos que orçamento total destinados às atividades de pesquisa. (DEWATRIPONT, et al., 2006, p. 8, tradução nossa)[2]

Este processo é chamado na literatura de Biblioteconomia e Ciência da Informação de "crise do periódico" (MUELLER, 2006) e afetou dramaticamente todo o sistema de comunicação científica então vigente, limitando as possibilidades das bibliotecas manterem suas coleções atualizadas. Como assinala Peek e Pomerantz (1998, tradução nossa): "O que ficou evidente foi que o relacionamento entre os editores e a comunidade acadêmica se desestabilizou"[3]. O Manifesto *Declaring Independence* (2001) chama este processo de "scientific communication crisis".

A partir da década de 1980 metodologias de avaliação de coleções de periódicos em bibliotecas tiveram que ser desenvolvidas para permitir selecionar quais assinaturas manter e quais excluir dos pacotes vendidos pelo publicadores comerciais. É importante assinalar para a compreensão deste processo, que até o surgimento da Internet, os publicadores desempenhavam um papel essencial e único no fluxo

---

2 "In the last 30 years, the prices of scientific journals have been steadily increasing. Between 1975 and 1995, they increased 200%-300% beyond inflation. This was accompanied by a fall in subscriptions both by individual researchers and by libraries whose budgets got squeezed. Indeed, journal prices far outpaced the evolution of library budgets, which did increase at a somewhat slower pace than total academic research budgets".

3 "What was apparent is that the relationship between publishing and the scholarly community has become destabilized".

internacional de ICT: o de disseminarem de forma ampla, através da venda de assinaturas e distribuição, em fascículos, o conhecimento científico, acumulado sob a forma das coleções de periódicos distribuídas nas bibliotecas especializadas e acadêmicas.

O próximo capítulo desta história, que define os contornos da problemática do livre acesso como acontece atualmente, é o surgimento da Internet e da WWW, em fins da década de 1980. O primeiro repositório digital de *preprints* surge em 1991, no laboratório de física de Los Alamos, Novo México, EUA, coordenado pelo físico Paul Ginsparg (1996). O primeiro periódico totalmente eletrônico, *The Online Journal of Clinical Trials,* surge em 1992 (PEEK; POMERANTZ, 1998).

A partir do início da década de 90 do século XX parcelas crescentes da comunidade acadêmica, incluindo aí associações de pesquisadores e de bibliotecas especializadas e acadêmicas, passam a buscar alternativas ao crescente custo de assinaturas imposto pelos publicadores comerciais. Com o surgimento da Internet a alternativa torna-se clara. Era possível publicar na Internet, a um custo mínimo, com um alcance mundial e com uma rapidez entre a submissão do artigo e sua publicação consideravelmente maior. A lógica das cobranças de assinaturas, começa a ser confrontada por vários setores da comunidade acadêmica, a lógica do livre acesso. Stevan Harnard (2001, p. 1), um dos mais combativos defensores do livre acesso, afirma que:

> Ao contrário dos autores de livros e artigos de revista, que escrevem para explorarem direitos ou por honorários, os autores de artigos de periódicos revisados por pares escrevem apenas pelo "impacto da pesquisa". Para ser citados e tomar parte na construção da pesquisa de outros pesquisadores, seus resultados têm de ser acessíveis aos seus usuários potenciais. Do ponto de vista dos autores, o acesso pago aos seus resultados é tão contraproducente como o acesso pago a anúncios comerciais [...],
>
> Os pesquisadores nunca se beneficiaram do fato de que as pessoas tinham que pagar para ter acesso aos seus documentos (na forma de assinaturas ou por acesso à versão *online*, na forma

> de licenças de acesso a portais ou via *pay-per-view*). Pelo contrário, estes obstáculos ao livre acesso representam barreiras impactantes para a trajetória dos pesquisadores, cujas carreiras dependem em grande parte da visibilidade e da aceitação de suas pesquisas.4

A relação visibilidade-acessibilidade-livre acesso, tão cara para a comunidade científica, também passa a ser crescentemente reconhecida:

> [...] artigos disponíveis *online* gratuitamente são muito mais citados. Para um impacto maior e um progresso científico mais rápido, os autores e os editores devem ter por objetivo tornar os resultados da pesquisas fáceis de serem acessados. (LAWRENCE, 2001, p. 521, tradução nossa)[5]

O movimento pelo livre acesso organiza-se politicamente e mais e mais desenvolve sua própria tecnologia. A Santa Fé Convetion (VAN DE SOMPEL; LAGOZE, 2000), uma reunião de gestores de repositórios de documentos científicos ocorrida em 1999, cria a Open Archives Initiative (OAI), com o objetivo de criar mecanismos tecnológicos para tornar interoperáveis os diferentes repositórios funcionando segundo a proposta Open Archives. Esses repositórios se proliferavam por diversos países e cobriam as diferentes áreas de conhecimento. No bojo desta iniciativa foi criado o padrão de metadados Dublin Core e o Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), para propiciar a coleta automática e o reuso de metadados de repositórios abertos (open archives) (WARNER, 2001).

---

4 "Unlike the authors of books and magazine articles, who write for royalty or fees, the authors of refereed journal articles write only for 'research impact'. To be cited and built on in the research of others, their findings have to be accessible to their potential users. From the authors' viewpoint, toll-gating access to their findings is as counterproductive as toll-gating access to commercial advertisements […] Researchers never benefited from the fact that people had to pay access tolls to read their papers (as subscriptions, and for the online version, site-licences or pay-per-view). On the contrary, those access barriers represent impact barriers for researchers, whose careers and standing depend largely on the visibility and uptake of their research."

5 "[…] articles freely available online are more highly cited. For greater impact and faster scientific progress, authors and publishers should aim to make research easy to access."

Outras iniciativas também surgem no sentido de prover o movimento pelo livre acesso com tecnologias de publicação eletrônica, de gestão de repositórios eletrônicos, como o Public Knowledge Project[6], iniciado em 1998 na Faculdade de Educação da Universidade de British Columbia, Canadá e o projeto EPrints, da Universidade de Southampton, Inglaterra, o mais usado *software* para gestão de repositórios que funcionem com a filosofia Open Access.

O próximo desdobramento do movimento pelo livre acesso é a sua mudança de qualidade no sentido de se tornar cada vez mais uma política pública, nos níveis institucional, nacional e internacional. Alguns marcos significativos desta trajetória são os seguintes:

- Lançamento do ArXiv, em 1991 - primeiro repositório eletrônico, no laboratório de física nuclear de Los Alamos, Novo México, EUA;
- Santa Fé Convention / Open Archives Initiative, em 1999. Santa Fé, Novo México, EUA - Propõe mecanismos tecnológicos de interoperabilidade entre esses repositórios eletrônicos para que o crescente número de repositórios que começa a se formar se torne um efetivo meio de comunicação científica;
- Scholarly Publishing & Academic Resources Coalition (SPARC) – uma associação mundial de bibliotecas especializadas, através do manifesto Declaring Independence[7], 2001: "Please join me in DECLARING INDEPENDENCE from publishers and journals that do not serve the research community.";
- Budapest Declaration[8], em 2001 - Evento do Open Society Institute;
- Primeira Instituição Acadêmica a adotar o Livre acesso a sua produção, School of Electronic and Computer Science, Univ. de Southampton, 2001.
- Declaração de Berlin[9], em 2003;
- Declaração de Bethesda[10], 2003;

6 http://pkp.sfu.ca/

7 http://www.arl.org/sparc/DI/

8 http://www.soros.org/openaccess

9 http://www.zim.mpg.de/openaccess-berlin/berlindeclaration.html

10 http://www.earlham.edu/~peters/fos/bethesda.htm

- WSIS 2003, Declaração de Princípios[11] (UNESCO), compromisso com livre acesso, item B3, 28;
- Resolução da Câmara dos Comuns[12], no Reino Unido, em 2004;
- Declaração de Salvador[13]: Commitment to Equity, durante o ICML 2005 - Ninth World Congress on Health Information and Libraries, Salvador, Brasil;
- Manifesto pelo Livre Acesso[14], Brasil, em 2005;
- Projeto de Lei n. 1.120[15], em 2007, Política de Livre acesso para o Brasil;
- Decisão dos pesquisadores da Univ. de Harvard[16] a favor do livre acesso, em 12 fev. 2008;

Dentro das políticas de livre acesso que vão se formulando por todo o mundo, os repositórios institucionais assumem um papel-chave. Longe de serem somente um aparato tecnológico, os repositórios institucionais se inserem como um instrumento dentro de uma política institucional, de determinada área de conhecimento ou comunidade acadêmica e, mesmo, nacional. Para se chegar ao livre acesso a informação científica de forma generalizada, o movimento pelo livre acesso propõe dois mecanismos, chamados de "vias": a via dourada, é uma orientação para que os periódicos científicos publiquem segundo a concepção do livre acesso. Esta via só é possível se os editores de periódicos aderirem à filosofia do livre acesso. De forma complementar a via dourada, é proposta também a via verde, que consiste justamente no depósito de trabalhos acadêmicos na rede de repositórios institucionais espalhadas crescentemente por todos os países do mundo, pelas mais diferentes instituições produtoras de conhecimento científico, tipicamente universidades, institutos de pesquisa e órgãos governamentais.

11 http://www.itu.int/wsis/docs/geneva/official/dop.html

12 http://www.publications.parliament.uk/pa/cm200304/cmselect/cmsctech/399/39903.htm

13 http://www.icml9.org

14 http://www.ibict.br/openaccess/arquivos/manifesto.htm

15 http://www.camara.gov.br/sileg/integras/461698.pdf

16 http://www.fas.harvard.edu/home/news-and-notices/news/press-releases/release-archive/releases-2008/scholarly-02122008.shtml

No nível de cada uma destas instituições, políticas locais, discutidas e aceitas pelas respectivas comunidades, regulam o depósito da produção científica dos pesquisadores da instituição. Estas políticas variam desde a simples sugestão de que os pesquisadores depositem cópias de livre acesso de seus trabalhos aceitos em periódicos ou congressos, passando por diferentes tipos de incentivos para que seja feito o autoarquivamento até a obrigatoriedade institucional de o fazerem. O depósito, dependendo da política, pode significar o livre acesso imediato ao trabalho científico, ou dentro de determinado prazo. Um número crescente de publicadores comerciais de periódicos começa também a aderir aos mecanismos da via verde.[17]

O autoarquivamento da produção acadêmica em repositórios institucionais de livre acesso também pode trazer novos e inusitados mecanismos de avaliação da ciência, que podem incentivar mais autores a realizarem o autoarquivamento (HARNARD, 2007). Alguns deles são: número de *downloads*, relação *downloads*/citação, índices de premiações e de apoios à pesquisa por pesquisadores etc.

Começa a se formar também toda uma infraestrutura complementar a iniciativa de arquivos abertos. Sem esta infraestrutura, a rede de repositórios de livre acesso será sempre vista como marginal aos fluxos internacionais de ICT. Fazem parte desta infraestrutura agregadores ou provedores de serviços, que permitem interoperabilidade e busca unificada sobre um conjunto de repositórios, propostas, diretrizes, modelos e mecanismos tecnológicos para garantir preservação digital (CCSDS, 2002) e repositórios altamente confiáveis. Um dos elementos-chave do que aqui é chamado de infraestrutura diz respeito a avaliação do impacto dos trabalhos depositados em repositórios. O Scientific Literature Digital Library and Search Engine (CiteSeer)[18], por exemplo, é um motor de busca que coleta e reúne metadados de um grande número de repositórios institucionais e temáticos, permitindo buscas integradas. Além disso, fornece para o universo de livre

17 Uma relação de periódicos científicos e de suas políticas relativas ao autoarquivamento e livre acesso por autores em repositórios institucionais pode ser encontrada em <http://romeo.eprints.org/>.

18 http://citeseerx.ist.psu.edu/

acesso importantes estatísticas e métricas de avaliação do impacto de pesquisas, similares à aquelas que no mundo dos publicadores comerciais são fornecidos pelo ISI (SHADBOLT et al., 2006).

Repositórios institucionais trazem agora para universidades e instituições de pesquisa a oportunidade de se fortalecerem institucionalmente a partir da visibilidade de sua produção acadêmica organizada e disponível, como um retrato fiel de sua instituição, a partir de seu repositório institucional. De um papel histórico e essencialmente passivo, como afirma Lynch (2002), essas instituições – cujo papel de produtoras de ICT era obscurecido pela relação direta que autores mantinham com os editores científicos – que tinham um papel no modelo UNISIST de meramente consumidoras de assinaturas de periódicos, passam a ter agora um papel destacado e cada vez mais essencial no cenário dos fluxos de ICT internacional. Os dirigentes destas instituições certamente considerarão isto um instrumento dos mais significativos para o desenvolvimento de políticas institucionais.

É um pouco este panorama que este livro quer trazer e aprofundar. Informação é um insumo essencial para o desenvolvimento da ciência e, consequentemente, para o desenvolvimento social e econômico de um país. O livre acesso à informação científica é um dos seus instrumentos, os repositórios institucionais, são atualmente um dos mais importantes desafios políticos e profissionais com que se deparam hoje os profissionais de informação no Brasil. O livre acesso à informação científica de boa qualidade, capaz de impulsionar a pesquisa brasileira, sempre foi um objetivo caro e especial de tantos sistemas de ICT implantados no Brasil desde o surgimento do CNPq e do IBBD na década de 1950. O desafio da implantação dos repositórios institucionais também retoma propostas metodológicas tão caras aos profissionais de informação, como o controle bibliográfico, a catalogação na fonte, o trabalho cooperativo. Este livro é dedicado aos profissionais de informação para fortalecê-los neste desafio.

REFERÊNCIAS

CCSDS - CONSULTATIVE COMMITTEE FOR SPACE DATA SYSTEM. Reference model for an Open Archival Information System (OAIS). Washington, 2002. Disponível em <http://public.ccsds.org/publications/archive/650x0b1.pdf >. Acesso em: 5 dez. 2009.

DECLARING INDEPENDENCE: a guide to creating community-controlled science journals. Washington, DC: Scholarly Pub. and Academic Resources Coalition; Chapel Hill, NC: Triangle Research Libraries Network, 2001. Disponível em: <http://www.arl.org/sparc/DI/>. Acesso em: 5 dez. 2009.

DEWATRIPONT, M. et al. **Study on the economic and technical evolution of scientific publications market in Europe**. Final Report – 2006. Brussels: DG-Research, European Commission, 2006. Disponível em: <http://dipot.ulb.ac.be:8080/dspace/bitstream/2013/9545/1/md-0005.pdf>. Acesso em: 09 out. 2009.

GINSPARG, P. Winners and losers in the global research village. In: CONFERENCE ON ELECTRONIC PUBLISHING IN SCIENCE, 1996, Paris. **Proceedings...** Disponível em <http://xxx.lanl.gov/blurb/pg96unesco.html>. Acesso em: 05 out. 2001.

HARNARD, S. Open access scientometrics and the UK research assessment exercise. In: ANNUAL MEETING OF THE INTERNATIONAL SOCIETY FOR SCIENTOMETRICS AND INFORMETRICS, 11., 2007, Madrid, Spain. **Proceedings...** Madrid, Spain : Cindoc, 2007. Disponível em : <http://eprints.ecs.soton.ac.uk/13804/>. Acesso em 2 dez. 2009.

______. The self-archiving initiative: nature web debates. 2001. **Nature web debates**. Disponível em: <http://www.nature.com/nature/debates/e-access/Articles/harnard.html>. Acesso em: 3 nov. 2001.

IBICT. **Manifesto brasileiro de apoio ao acesso livre à informação científica**. Brasília, 2005.

LAWRENCE, S. Free online availability substantially increases a paper's Impact. **Nature web debates**. Disponível em: <http://www.nature.com/nature/debates/e-access/Articles/lawrence.html>. Acesso em: 10 jun. 2001.

LYNCH, C. A. Institutional repositories: essential infrastructure for scholarship in the digital age. **ARL**, n. 226, p. 1-7, Feb. 2003, Disponível em: <http://www.arl.org/resources/pubs/br/br226/br226ir.shtml>. Acesso em: 31 jul. 2009.

MUELLER, S. P. M. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**, v. 35, n. 2, ago. 2006 . Disponível em <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652006000200004&lng=en&nrm=iso>. Acesso em: 27 nov. 2009.

PEEK, R. P.; POMERANTZ, J. P. Electronic Scholarly Journal Publishing.

Annual Review of Information Science and Technology, Medford, v. 33, p. 321-356, 1998,. Disponível em <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.98.5541&rep=rep1&type=pdf>. Acesso em: 27 nov. 2009. Preprint.

RIEGER, O. Y. Opening up institutional repositories: social construction of innovation in scholarly communication. **Journal of Electronic Publishing,** v. 11, n. 3, Fall 2008. Disponível em <http://dx.doi.org/10.3998/3336451.0011.301>. Acesso em: 6 ago. 2009.

SHADBOLT, N. et al. The open research web: a preview of the optimal and the inevitable. 2006. Disponível em <http://cogprints.org/4841/2/shad-bch.pdf>. Acesso em: 24 ago. 2009.

UNISIST - UNITED NATIONS INTERNATIONAL SCIENTIFIC INFORMATION SYSTEM. **Study Report on the Feasibility of a World Science Information System, by the United Nations Educational, Scientific and Cultural Organization and the International Council of Scientific Unions**. Paris: UNESCO, 1971.

VAN DE SOMPEL, Herbert; LAGOZE, Carl. The Santa Fe Convention of the Open Archives Initiative. **Dlib Magazine**, v. 6, n. 2, Feb. 2000. Disponível em: <http://www.dlib.org/dlib/february00/vandesompel-oai/vandesompel-oai.html>. Acesso em:.

WARNER, Simeon. Exposing and harvesting metadata using the OAI Metadata Harvesting Protocol: a tutorial**. High Energy Physics Libraries Webzine**, v. 4, June, 2001. Disponível em: <http://library.cern.ch/HEPLW/4/papers3>. Acesso em: 18 nov. 2001.

Software livres para repositórios institucionais: alguns subsídios para a seleção

Luis Fernando Sayão
Carlos Henrique Marcondes

INTRODUÇÃO

Por todo o mundo, as universidades e os centros de pesquisa estão, de uma forma muito intensa, planejando, implementando ou operando repositórios institucionais. Esse movimento coletivo é parte do esforço das comunidades acadêmicas de explorar as possibilidades oferecidas pelas tecnologias de informação e comunicação para criar serviços de informação inovadores, através dos quais as novas mídias digitais possam potencializar o ensino, a pesquisa e a comunicação científica.

Numa primeira aproximação, a implantação de um repositório institucional é o reconhecimento de que as atividades intelectuais e acadêmicas das instituições de pesquisa e ensino estão crescentemente representadas, documentadas e compartilhadas em formato digital; e que uma das principais responsabilidades dessas instituições de conhecimento é exercitar a custódia sobre esses conteúdos no sentido de torná-los disponíveis para o acesso e para preservá-los por longo prazo. Os repositórios institucionais são o meio pelo qual essas instituições estão endereçando esta responsabilidade para os membros das suas comunidades e para o público em geral (LYNCH, 2003).

No contexto do presente estudo, podemos pensar no repositório institucional como uma base de dados na Web na qual uma instituição de pesquisa deposita sistematicamente sua produção acadêmica e a disponibiliza de forma ampla para as comunidades interessadas. Sobre essa base de dados é oferecido um conjunto de serviços voltados para

a gestão e para a disseminação de informações em formato digital. Esses serviços incluem captura, armazenamento, tratamento técnico, organização, preservação e entrega de conteúdos digitais de toda a natureza – texto, imagens, vídeo, áudio, apresentações, programas de computador, *datasets* etc.

Como o seu próprio nome expressa, via de regra, os repositórios institucionais são serviços vinculados organicamente aos seus ambientes institucionais; são cumulativos e persistentes e, portanto, têm compromissos fortes com a formação da memória digital acadêmica, com a preservação de longo prazo de materiais de valor contínuo e com os movimentos de livre acesso. Para tal, esses repositórios têm como base tecnológica sistemas abertos e interoperáveis e aderência aos padrões das áreas de Biblioteconomia, Ciência da Informação e Tecnologia da Informação.

A criação de repositórios institucionais compreende um grande número de atividades que ensejam aspectos políticos, legais, educacionais, culturais e alguns componentes técnicos importantes. O encaminhamento correto desses vários aspectos e de suas interrelações é que vai determinar o perfil do repositório e a sua aproximação aos objetivos fixados pela organização e, por fim, o sucesso do empreendimento.

As comunidades interessadas, que hoje já extrapolam o domínio acadêmico, têm lançado mão dos repositórios institucionais para um espectro extraordinário de funções que variam significantemente de repositório para repositório. Muitas organizações utilizam repositórios institucionais para aplicações e usos inéditos, por exemplo: gestão de atividades de pesquisa, veiculação de publicações eletrônicas, armazenamento de materiais de aprendizagem, gestão de dados de pesquisa, curadoria de materiais digitais, gestão de conhecimento, exposição virtual, para citar alguns.

Para que seja factível desenvolver repositórios institucionais adequados a políticas e a modelos específicos e que, ao mesmo tempo, acomodem a pluralidade de interesses das diversas comunidades, é necessário adotar plataformas de *software* versáteis, com capacidade

de expansão e de integração a outros programas que possam apoiar o atendimento às demandas atuais e futuras.

A boa notícia é que há uma extraordinária oferta de *software* para implementação de repositórios digitais e, surpreendentemente, os mais sofisticados, técnica e funcionalmente, são livres e de código aberto. Esse fato coloca um desafio importante para quem deseja desenvolver um repositório: selecionar a plataforma de *software* mais adequada aos requisitos técnicos, funcionais e gerenciais estabelecidos como perfil para o repositório.

Para a avaliação de repositórios institucionais – considerando-os como um conjunto de serviços de informação que se materializam por meio de um site na Web – é necessário a concorrência de critérios de avaliação para serviços e sistemas de informação e de critérios mais próximos à avaliação de Websites, como os aspectos de valoração dos contextos digitais – por exemplo: a organização da informação nesses contextos e a interação dos usuários com eles.

Na definição de critérios gerais para avaliação de bibliotecas digitais, Céspedes (2006) confirma que os modelos de avaliação devem se desenrolar segundo três dimensões: 1) biblioteconômica – onde se consideram os componentes que contextualizam os sistemas de informação: usuários, serviços e produtos de informação, elementos de organização e representação de conteúdos, recursos de recuperação e busca, e outros; 2) tecnológica – dimensão que parte da concepção da biblioteca como um site na Web e da análise de *hardware, software* e infraestrutura de rede disponíveis para a sua implementação; 3) interação usuário-sistema – cuja ênfase está no design das interfaces, nas funcionalidades do sistema e nas demais facilidades através das quais os usuários interagem com os conteúdos digitais gerenciados pelo sistema.

Adicionalmente, é necessário ainda incluir uma dimensão voltada para a gestão e administração do repositório que incluem, por exemplo, a gestão de direitos e de acesso, da preservação digital, das coleções, da segurança da informação e ainda os instrumentos de gestão voltados para o usuário final.

Um fator determinante na viabilização dessas dimensões é a qualidade da plataforma de *software* e a sua adequabilidade aos requisitos funcionais sobre a qual o repositório irá operar. Portanto, o desafio de selecionar uma plataforma de *software* como parte do ciclo de planejamento de repositórios institucionais torna-se uma fase crítica para o sucesso do empreendimento.

Circunscrito a esta perspectiva, o presente trabalho volta a atenção para a definição de uma matriz de critérios que sirvam de subsídios no processo de avaliação e seleção desses *softwares,* com ênfase nos requisitos técnicos e funcionais.

O conjunto de critérios sugerido considera as funcionalidades necessárias à implantação, à gestão e à operação do repositório e ao uso dos conteúdos digitais. À guisa de ilustração, as características técnicas de alguns dos *softwares* mais usados e/ou mais conhecidos no país são explicitadas. São eles: DSpace, Eprints, Greeenstone, Nou-Rau e Fedora.

ALGUMAS CONSIDERAÇÕES SOBRE *SOFTWARE* PARA REPOSITÓRIOS DIGITAIS

A concepção de repositórios digitais como sendo um simples armazenador estático de informação digital com capacidade de recuperação foi rapidamente ultrapassada. A ideia original deslocou-se para um conceito mais sofisticado de sistema de informação que incorpora a facilidade da comunicação, da colaboração e de outras formas de interação dinâmica entre usuários de um vasto universo.

Na primeira geração de bibliotecas digitais, quando as coleções eram pequenas e tinham um caráter essencialmente experimental, uma grande variedade de programas não especificamente voltados para aplicações de repositórios digitais foram utilizados, como os gerenciadores de bases de dados bibliográficos – por exemplo, o Micro-ISIS –, e ainda os *softwares* gerenciadores de banco de dados mais genéricos, como o MS Access. Nestes primeiros momentos, uma biblioteca digital, não era nada mais do que uma coleção de recursos eletrônicos colocados disponíveis numa página Web ou em um CD-ROM.

A situação, entretanto, evoluiu rapidamente: hoje os principais projetos de bibliotecas e de repositórios digitais estão colocando *on-line* estoques massivos de recursos informacionais em formato digital. Algumas dessas coleções incluem milhões de objetos e estão sendo planejadas para gerenciarem um número astronômico de informações de toda a natureza num futuro bem próximo.

Com a evolução dos repositórios em termos de volume, diversidade de materiais em formato digital – que se multiplicam cotidianamente – e de funcionalidades, tendo ainda como perspectiva imediata a interoperabilidade, se torna crítico que a plataforma tecnológica subjacente aos repositórios digitais seja capaz de apoiar com desempenho e confiabilidade a trajetória de complexidade ascendente desses projetos.

Ainda é necessário considerar que o rápido crescimento de tipos variados de repositórios digitais é um dos primeiros desdobramentos do desenvolvimento baseado em padrões abertos de arquitetura e de *software*. Sem esses padrões, teríamos poucas e caras ilhas de acervos digitais que só poderiam ser acessadas via sistemas especiais projetados para cada uma delas.

Entretanto, a tecnologia que envolve os repositórios digitais guarda uma peculiaridade marcante e positiva que deve ser explorada. No ciclo de automatização anterior, cujo esforço estava centrado no desenvolvimento de pacotes de *software* para automatização de bibliotecas, a oferta de sistemas estava – e ainda está – dominada por pacotes comerciais. São raros os programas disponíveis livremente com qualidade profissional. Em contraste, os principais programas para criação de repositórios digitais são distribuídos livremente para uso e para desenvolvimentos posteriores. Este fato é uma consequencia de uma característica quase que comum no desenvolvimento desses programas: tipicamente eles foram resultados de projetos de pesquisa que juntaram universidade, agências governamentais, organizações internacionais e, em alguns casos, empresas e organizações não governamentais.

Como resultado dessa origem "acadêmica", há uma oferta considerável de *software* de códigos abertos e distribuídos livremente voltados para a implementação de repositórios digitais de toda natureza. Esses

programas apresentam características sofisticadas, elevado grau de qualidade e, sobretudo, conformidades aos principais padrões da área, estabelecendo, dessa forma, as bases para a integração e a interoperabilidade. A oferta é diversificada, vai desde *softwares* prontos para instalação e uso imediato, até pacotes mais próximos a *kits* de ferramentas que para serem utilizados demandam recursos de programação para o desenvolvimento de aplicações e de interfaces.

ESCOLHA DA PLATAFORMA DE *SOFTWARE*

Uma vez definida a necessidade de se criar um repositório institucional, fixadas as políticas de gestão e de uso do repositório, e estabelecido o elenco de serviços que serão oferecidos a uma comunidade definida de usuários, é chegado o momento de examinar as opções de *software* disponíveis – considerando que a opção da instituição não é o desenvolvimento de uma solução própria. Para tal, é necessário compor um conjunto de requisitos que refletirão as diretrizes e as políticas gerais adotadas pelo repositório, que poderemos chamar de perfil do repositório. Esses requisitos vão estabelecer uma métrica para avaliação dos pacotes de *software* disponíveis.

Para elaborar o conjunto de requisitos e para avaliar as alternativas de *software* é sempre recomendável formar um comitê integrado por membros da administração da biblioteca, por profissionais de tecnologia da informação e representantes dos vários segmentos de usuários potenciais do repositório. Cada membro do comitê irá contribuir com seus conhecimentos e sua experiência prática sobre como o sistema deve operar e sobre as características desejáveis do sistema, em termos de funções e serviço e das características técnicas subjacentes – sistema operacional, banco de dados, mecanismos de busca etc. (BARTON; WATERS, 2004).

CARACTERÍSTICAS GERAIS

Os itens relacionados a seguir são características gerais empregadas

na avaliação da maioria dos pacotes de *software*, não importando o seu escopo de aplicação e suas especificidades. Não obstante o seu caráter geral, essas características devem ser incorporadas às metodologias de avaliação dos programas empregados para criação de repositórios institucionais. (MARILL; LUCZAK, 2009)

- Escalabilidade – avalia a capacidade de crescimento do sistema por meio de adição de mais recursos (CPU, RAM etc.) para gerenciar coleções crescentemente maiores e diversificadas de objetos digitais. Tendo em vista a perspectiva de crescimento do repositório, este pode ser um critério crítico;
- Extensibilidade – mede a capacidade do programa de integrar ferramentas externas no sentido de estender as funcionalidades do repositório;
- Facilidade de implantação – avalia o grau de simplicidade no processo de instalação e de configuração; avalia a facilidade de integração com outros programas necessários ao funcionamento do repositório;
- Plataforma computacional – identifica os componentes necessários ao sistema: sistema operacional (por exemplo: Windows, Unix/Linux etc.), servidores Web (por exemplo: Apache), *software* gerenciador de banco de dados (por exemplo: MySQL, Postgres, Oracle, SQL Server etc.); identifica a exigência de outros *software* clientes que não seja o *web browser*;
- Implantações de sucesso – estima o número de usuários e organizações satisfeitos com o programa; analisa o histórico de implementações e a vitalidade e a sustentabilidade do pacote;
- Suporte do sistema – avalia a capacidade de resposta dos desenvolvedores e/ou da comunidade de usuários (para *software* livre) a problemas técnicos ou de outra natureza colocados;
- Base de conhecimento das comunidades envolvidas – estima a base de conhecimento e o nível de atividade das comunidades envolvidas no desenvolvimento e uso do programa; avalia a qualidade e a completeza das informações presentes no *site* do produtor e dos canais de comunicação – e-mail, fóruns, eventos, newsletters etc.;

- Estabilidade da organização de desenvolvimento – estima o grau de confiança na organização responsável pelo desenvolvimento do programa: histórico, tradição, tempo de existência, sustentabilidade econômica; inserção e relacionamentos com outras organizações etc.;
- Perspectivas para o futuro – avalia a capacidade de evolução e de incorporação de inovações;
- Limites do sistema – avalia os limites do programa: volume de dados, número de coleções, de registros, de bases de dados etc. que o *software* consegue gerenciar;
- Documentação disponível/cursos/publicações – estima a qualidade e o volume de informações sobre o pacote elaborado pela instituição produtora; estima as informações disponíveis sobre o pacote provenientes de outras fontes – livros comerciais, tutorias, cursos, artigos etc.

Outros itens podem compor o conjunto de critérios acima, tais como: mantenabilidade, confiabilidade, equipe de programação e operação necessária etc.

DISTRIBUIÇÃO DO PACOTE DE SOFTWARE

Os pacotes de *software* para criação de repositórios digitais estão disponíveis por meio de diferentes modelos de licença e de distribuição. Os principais são os seguintes:

- *Software* proprietários – o cliente paga por uma licença de uso do *software* e, via de regra, subscreve as atualizações do programa e contrata assistência técnica; de forma geral, é oferecido ao cliente graus variados de customização do *software*, porém o código fonte pertence à organização que criou e que mantém o pacote.
- *Software* de código aberto – tipicamente o cliente baixa gratuitamente o *software* da Internet e o instala por conta própria ou assistido por consultoria contratada; a equipe de tecnologia da informação pode examinar, customizar e melhorar o código fonte; uma organização central gerencia o código fonte que,

entretanto, está aberto para mudanças e aperfeiçoamentos implementados pela comunidade de desenvolvedores. Existem várias modalidades para este tipo de licença, por exemplo: BSD, GNU Public Licence, Mozzila.

É necessário estar atento aos custos não explícitos de ambos os modelos. Pode ser bastante útil consultar os responsáveis pela implementação de repositórios em outras organizações que utilizaram os pacotes que estão sendo avaliados (BARTON; WATERS, 2004).

ENQUADRAMENTO INFRAESTRUTURA, ORGANIZACIONAL E ESTRATÉGICO

Não obstante os modelos de avaliação estarem focados nas qualidades técnicas e funcionais dos pacotes, a decisão sobre a plataforma de *software* que será utilizada não deve estar baseada unicamente nessas características. É necessário considerar o enquadramento do *software* às diretrizes e políticas organizacionais e estratégicas da instituição, bem como o seu ajustamento à infraestrutura disponibilizada por ela para o projeto (POWEL, 2005).

Infraestrutura técnica, gerencial e metodológica

Define quais são os requisitos mínimos de *hardware*, de *software* e de infraestrutura de rede necessários para a instalação, plena operação e manutenção do *software*; deve incluir a disponibilidade interna ou contratada de capacidade de desenvolvimento e programação, o que é muitas vezes exigido pelos pacotes, e a disponibilidade de recursos – humanos, financeiros e metodológicos – continuados para a operação e gerência do repositório. Deve considerar:

- *Hardware* – disponibilidade dos requisitos mínimos de *hardware* – máquinas servidoras; memória, capacidade de armazenamento, processador etc.; infraestrutura de rede: roteadores, largura de banda etc.
- *Software* – disponibilidade dos *software* necessários: ambiente ope-

racional; servidor Web, banco de dados, linguagens, ferramentas de indexação e de busca etc.

- Expertise técnica gerencial e metodológica – disponibilidade interna ou contratada de conhecimento e de experiência para a instalação e operação do *software* e para a gestão do sistema como um todo.
- Sustentabilidade – disponibilidade de fundos para planejar, implantar, operar e manter o repositório; deve ficar claro que os custos das plataformas de *software* e de *hardware* são geralmente previsíveis, entretanto os custos do planejamento total, da implementação, do treinamento da equipe, de eventuais obras e instalações, consultorias externas e de operação do sistema são dependentes do alcance e grau de sofisticação do projeto do repositório.

Enquadramento à política de TI da instituição

Além de avaliar se a instituição dispõe da infraestrutura exigida para instalação do programa e para operação e gestão do repositório, é importante considerar se a infraestrutura técnica necessária está alinhada à política e à estratégia de tecnologia de informação da organização e ao seu plano de negócios. Não há possibilidade de sustentabilidade do repositório sem um razoável enquadramento institucional, posto que há sempre o risco de se criar um gueto tecnológico, excluído dos planos de desenvolvimento e manutenção e das estratégias futuras da instituição.

SERVIÇO E FUNCIONALIDADES

Os requisitos definidos pela instituição em termos de serviços e de funções para o repositório estão geralmente formalizados no documento de especificação de requisitos funcionais. O documento serve de base para a avaliação do grau de atendimento dos pacotes de *software* aos interesses da instituição e ao perfil definido para o repositório.

As tecnologias da informação e de comunicação trazem possibilidades de criação de novos serviços e de formas inéditas de entrega da informação em formato digital, entretanto, os repositórios digitais ainda têm como base muitos dos serviços tradicionais de biblioteca, como busca e disseminação seletiva de informação. Alguns indicadores gerais para os serviços de um repositório que são dependentes do *software* são os seguintes:

- Diversidade de serviços – estima a diversidade de serviços e produtos de informação que o *software* disponibiliza de forma nativa ou permite implementar com seus próprios recursos ou integrando-se a outros programas.
- Interatividade – indica o grau de interatividade em tempo real do usuário com o sistema que é possível implementar com o pacote.
- Tempo de resposta – mede o espaço de tempo entre uma requisição e a resposta satisfatória do sistema.

Interfaces do Sistema

O conjunto de serviços é a face mais visível para o usuário dos sistemas de informação seja ele profissional de informação, exercendo trabalhos técnicos, ou usuário final depositando e acessando conteúdos digitais. Porém, a visão das diversas categorias de usuários sobre os serviços e funcionalidades do repositório é intermediada pelas diversas interfaces do sistema. No presente escopo, podemos limitar a análise a dois tipos de interfaces: interfaces profissionais, voltadas para entrada de dados, tratamento técnico e gestão do repositório, e interfaces do usuário final voltadas para autopublicação, busca, *browsing* e acesso aos conteúdos digitais.

Geralmente os *software* oferecem interfaces *default* que podem ser personalizadas, entretanto, em algumas situações, as interfaces devem ser desenvolvidas pelo equipe de TI da instituição. Os seguintes itens são tomados como critérios de avaliação geral (as interfaces específicas serão analisadas em outro momento):

- Interface *default* – verifica se o *software* oferece interfaces prontas

para o uso, sem necessidade de programação.

- Padrões de interface – verifica se as interfaces estão em conformidade com os padrões (abertos) apropriados.
- Customização – verifica se as interfaces podem ser adaptadas às exigências de aparência da instituição – *lay-out,* cores, logomarcas etc.
- Ergonomia e usabilidade – avalia se a interface é ergonômica, facilitando a usabilidade da biblioteca em termos de navegação, ajuda *on-line,* nível de experiência do usuário etc.
- Acessibilidade – avalia se a interface implementa os padrões básicos de acessibilidade;
- Multilíngue – verifica se a interface é multilingue e em que idiomas é possível interagir com o sistema; verifica se é possível adicionar outros idiomas e as dificuldades para tal.

Incorporação dos objetos no repositório

Circunscreve os requisitos oferecidos pelo programa para a inclusão de objetos digitais no repositório: interfaces; formatos aceitos; importação de dados; modelos de objetos complexos; autoarquivamento e *workflow*.

- Interface de entrada de dados – avalia as interfaces de entradas de dados disponíveis e suas qualidades: facilidade de edição, acesso à lista de autoridades e a tabelas, ajuda *on-line,* adaptação ao tipo de material etc.
- *Software* cliente – avalia a necessidade e a disponibilidade de *software* cliente para a entrada de dados que não sejam os navegadores (*browsers*) padrão.
- Formatos aceitos – cotidianamente testemunhamos o surgimento de novos formatos de objetos digitais proporcionado pelo avanço da tecnologia da informação, portanto é necessário que os *softwares* de repositório digital sejam capazes de dar o tratamento adequado a essa diversidade de formatos. Desta forma este item avalia o conjunto de formatos de objetos digitais – PDF, HTML, MP3, GIF, por exemplo – aceitos por *default* pelo programa; avalia também as facilidades oferecidas pelo programa para processa-

mento de outros formatos, por exemplo, através do uso de *plugins* ou por conversão para formatos permitidos.

- Importação de dados de outras fontes – verifica se o programa tem capacidade de importar dados de outras fontes ou sistemas – por exemplo: bases de dados, OPACs – Online Public Access Catalogs, repositórios digitais; caso esta operação seja permitida, verifica que padrões e protocolos são usados para isto, por exemplo, base de dados Micro-ISIS (ISO2709), formato MARC 21, padrão Metadata Encoding Transmission Standard (METS) etc. (ver também item 3.5.5)
- Modelo de objetos complexos – uma parte significativa dos repositórios acadêmicos é desenvolvida em torno da ideia de objetos simples, que podem apresentar múltiplas versões e/ou vários formatos, mas que, no entanto, apresentam um único registro de metadados para este objeto. Os metadados criam uma vinculação entre as diversas versões/formatos. Porém, em algumas situações, o conteúdo é ingressado no repositório como um pacote, formando um maço complexo de recursos interligados, como é o caso dos objetos educacionais, páginas *web* e objetos multimídiaticos. Cada um dos itens do pacote pode ter registros de metadados separados associados a eles e devem ser vinculados por "empacotadores" como o determinado pelo padrão METS que facilita a maneira como esses tipos de objetos podem ser depositados, gerenciados e recuperados (POWELL, 2005; SAYÃO; MARCONDES, 2008). Este requisito verifica se o *software* implementa um modelo de objeto digital que permita a incorporação de versões, instâncias, etc; verifica se o programa oferece suporte aos padrões de empacotamento de metadados mais comuns, o METS e o MPEG-21;
- Autoarquivamento/autosubmissão/revisão/aprovação – avalia a disponibilidade da função de autoarquivamento, i.e, módulo que permite o autor submeter o seu próprio trabalho, fornecendo metadados e fazendo *upload* do seu conteúdo; avalia também a existência de módulo que viabiliza a retenção do trabalho para revisão, crítica e edição de seus metadados até que ele seja libe-

rado pelo editor ou administrador da biblioteca para publicação.

- Objetos de dados distribuídos – avalia se o conteúdo pode ser referenciado não somente no ambiente local, mas em qualquer lugar na *web*.
- Fluxo de trabalho (*workflow*) – está relacionado à disponibilidade de funções de *workflow* no controle do processo de entrada de dados – por exemplo, aprovação de conteúdos – e outros processos gerenciais.
- Armazenamento – avalia as formas de armazenamento dos conteúdos e metadados submetidos ao repositório: conversão para XML, PDF etc., armazenamento distribuído, possibilidade de armazenamento em mídias portáveis (por exemplo: CD-ROM ou DVD)

Organização da Informação

A forma com que o *software* apóia a organização da informação principalmente para a navegação (*browsing*) é, para alguns tipos de repositórios, de grande importância, posto que reflete mais espontaneamente a natureza da instituição. Por exemplo, um arquivo digital precisa que os seus objetos estejam organicamente agrupados em fundos, dossiês, processos; já para um repositório institucional de um centro de pesquisa talvez seja interessante uma organização por comunidades – por exemplo, laboratórios, departamentos, projetos, etc.

- Formas de organização – avalia se o programa tem funcionalidades para organizar a informação conforme as necessidades da instituição.

Indexação

Geralmente, cada *software* emprega algoritmos próprios para indexar as informações que descrevem os conteúdos com a finalidade de descoberta de recursos e navegação. Para isto eles utilizam os metadados descritivos que foram assinalados por quem entrou com os dados e/ou extraem ou criam metadados automaticamente a partir

dos conteúdos que estão sendo incorporados. De forma diferente dos programas de automação de bibliotecas, é muito comum que os programas voltados para repositórios digitais utilizem ferramentas de indexação independentes, desenvolvidos por outros produtores; cada uma dessas ferramentas utiliza algoritmos e métodos de indexação distintos; é comum também que o *software* de repositório ofereça a possibilidade do gestor escolher a ferramenta e o método de indexação mais conveniente.

- Métodos de indexação – identifica qual é o método de indexação utilizado pelo pacote; verifica se o método é nativo ou é implementado por *software* independente; verifica se o *software* de indexação é livre ou proprietário; verifica os controles e ajustes possíveis.
- Texto completo – verifica se o *software* possui módulo de indexação do texto completo; verifica o método utilizado.
- Adequação da indexação - avalia se o método empregado é adequado para as exigências de recuperação do repositório;
- Criação de índices – verifica .se o *software* cria índices (por exemplo, por autor e assunto) para *browsing*.
- Vocabulários controlados – avalia se o *software* tem módulo ou possibilidade de integração a ferramentas de apoio à criação de vocabulários controlados, lista de autoridades, etc.

Recuperação da Informação

A recuperação é uma decorrência imediata da capacidade de indexação do sistema e da qualidade dos metadados assinalados por quem entra com os dados ou dos indexadores extraídos automaticamente pelo programa. Devem ser avaliadas as possibilidades de busca oferecidas por *default* pelo programa e as ferramentas disponíveis para customização.

- Modalidade de buscas disponíveis – avalia as modalidades de busca oferecidas pelo programa: busca simples, avançada, profissional etc.; possibilidade de consulta aos índices para formulação de busca etc.
- *Browsing* – Avalia se o programa oferece possibilidade de navegação pelos índices.

- Disseminação de informação – Avalia se o programa tem disponível, por *default*, outras funcionalidades voltadas para a disseminação de informações, por exemplo: DSI, RSS, alertas etc.
- Apresentação dos resultados – identifica as opções de exibição dos resultados da busca: ordenação, tipos de formatos (curtos, longos, customizados etc.) e outros.
- Acesso aos conteúdos – avalia as opções de acesso aos conteúdos: *downloading, streaming*, OpenURL etc.
- Interface de consulta – avalia a interface de consulta segundo os parâmetros de: customização, usabilidade, acessibilidade, níveis de experiência do usuário, etc.

Interatividade entre usuários

Uma característica que distingue os repositórios digitais de outras bases de dados e a disponibilidade ferramentas que viabilizem a interação e a troca de idéias entre usuários mediadas por moderador ou não. A possibilidade de se submeter comentários sobre determinada publicação é uma das formas mais características de interação entre pesquisadores, entretanto outras modalidades podem estar presentes.

- Comentários – avalia se o *software* disponibiliza facilidades para submissão de comentários ou outras modalidades de interação; avalia as possibilidades de moderação.

CONFORMIDADES COM PADRÕES

Metadados e Esquemas de Metadados

Segundo a National Information Standards Organization (NISO)[1], metadados é informação estruturada que descreve, explica, localiza, ou ainda possibilita que um recurso informacional seja fácil de recuperar, usar ou gerenciar. O termo metadados frequentemente designa dados sobre dados, ou informação sobre informação. Os metadados estão

1 <http://www.niso.org/home/>

categorizados em metadados descritivos, voltados para facilitar a descoberta, a identificação, a compreensão e a seleção de recursos; metadados administrativos que facilitam a gestão, o acesso e a preservação dos recursos digitais; e os metadados estruturais que documentam a estrutura dos objetos e os relacionamentos entre objetos digitais. Os metadados ainda desempenham um papel importante no domínio da interoperabilidade. Existem vários esquemas de metadados com graus diferenciados de especificidade, porém o mais importante deles é o Dublin Core, considerado a língua franca para representação de recursos na *web*. O esquema é composto por 15 elementos que foram projetados para serem de simples compreensão e de fácil aplicação pelo próprio autor. Entretanto o esquema permite que diferentes comunidades façam adaptações e o customizem, adicionando elementos, redefinindo e ampliando a semântica deles, o que pode ser facilitado pelo uso de qualificadores, formando o que se chama de "perfil de aplicação". Dessa forma é necessário avaliar:

- Esquemas de metadados aceitos – verifica quais são os esquemas padronizados de metadados aceitos pelo pacote.
- Dublin Core – verifica se o programa aceita o esquema Dublin Core; caso positivo verifica se ele permite a extensão do esquema através do uso de qualificadores e/ou se é possível adicionar elementos e criar perfis de aplicação adequados à comunidade usuária.
- Criação/edição de metadados – avalia se o programa dispõe de facilidade para a criação e para a edição de esquemas de metadados.

Preservação Digital

Uma das mais importantes motivações para a criação dos repositórios institucionais é assegurar que os materiais digitais de pesquisa permaneçam disponíveis e acessíveis por longo prazo, contribuindo para a construção e preservação da memória acadêmica das instituições de pesquisa e ensino. Nessa direção, espera-se que os repositórios digitais disponham de metodologias e ferramentas que mantenham íntegros estes estoques por longo prazo. A norma mais importante da área é o Open Archival Information System (OAIS), um modelo conceitual

desenvolvido pela NASA e tornado uma norma ISO em 2002; o OAIS estabelece um modelo de informação e um modelo funcional, ou seja, descreve as informações e as funções necessárias para a preservação de longo prazo no âmbito de um repositório digital (SAYÃO, 2006). Portanto, os itens que devem ser considerados são:

- Norma ISO/OAIS – avalia o grau de aderência – parcial ou total – do *software* à norma OAIS e a sua capacidade de implementar os modelos de informação e funcional conforme especificada pelo modelo OAIS.
- Outras metodologias – avalia se o *software* apóia ou aplica outras metodologias – cumulativamente ou não com o OAIS –, como por exemplo, a conversão das informações para formatos padronizados, como o XML ou texto puro;
- Migração – avalia a disponibilidade de ferramentas de apoio à gestão do processo de migração;
- Outras estratégias de preservação digital – avalia a disponibilidade de aplicação de alguma outra estratégia de preservação.

Interoperabilidade

O conceito pleno de biblioteca digital pressupõe a troca de informação entre bibliotecas e a possibilidade de busca em repositórios heterogêneos administrados por diferentes instituições organizadas em diferentes níveis de federação. Para o usuário isto deve acontecer de forma transparente e através de uma única interface. Um fator de fundamental importância para a interoperabilidade são os protocolos e padrões abertos que devem ser aplicados em todas as instâncias. No domínio bibliográfico, atualmente, dois protocolos, de enfoques distintos, exercem um papel chave, o OAI-PMH e o Z39.50. Portanto é necessário avaliar a aderência a esses padrões:

- Protocolo Open Archives – verifica se o pacote implementa o protocolo OAI/PMH;
- Protocolo Z39.50 – verifica se o pacote implementa o protocolo Z39.50.

Entretanto, vários outros protocolos mais gerais da área de T.I. têm sido usados cada vez mais no mundo das bibliotecas digitais (SAYÃO; MARCONDES, 2008). Por exemplo:

- Outros protocolos de interoperabilidade – verifica se o pacote implementa outros padrões de interoperabilidade, por exemplo: SOAP e REST Web Service, SRW (Search/Retrieval Web Service), SRU (Search Retrieval via URL)

Identificadores/Nomes Persistentes

O objetivo dos sistemas de identificadores persistentes é garantir que os recursos informacionais tenham uma identificação – diferentemente do URL - única, global, e independente de tecnologia e do endereço do servidor onde o recurso está armazenado. Isso assegura que o recurso possa se materializar sempre que o *link* correspondente for acionado, contribuindo para o estabelecimento de sistemas interoperáveis e para a preservação digital (SAYÃO, 2007). Nesse sentido, um dos itens importantes na avaliação de um pacote de *software* de repositório digital é saber se ele implementa algum dos sistemas de identificadores persistentes já estabelecidos, como por exemplo o Persistent URL (PURL), o Data Object Identifier (DOI), o Uniform Resource Name (URN) ou o Handle System, ou se implementa alguma estratégia específica.

- Nome padronizado – avalia se o *software* implementa algum sistema padronizado de identificação permanente: URN, DOI, Handle System;
- Esquema específico – avalia se o *software* implementa alguma estratégia específica de identificação dos seus conteúdos e se ela está em conformidade com os padrões apropriados; por exemplo, PDI implementado pelo pacote Fedora.

Importação/exportação de dados

Para os repositórios digitais é de grande importância a possibilidade de importar dados de outros sistemas, por exemplo, da base de dados

da biblioteca, e vice-versa. Esta é uma funcionalidade crítica para o povoamento inicial dos repositórios.

- Importação/exportação – verifica se o *software* é capaz de intercambiar dados com outros sistemas por meio de padrões/protocolos abertos.
- Padrões – verifica os padrões abertos aceitos pelo *software* para importação/exportação de dados: METS, XML, ISO2709, MPEG21, MARC, etc.

GESTÃO DO REPOSITÓRIO E SEGURANÇA DA INFORMAÇÃO

Para serem gerenciados, os repositórios institucionais necessitam que os pacotes de *software* ofereçam funcionalidades que auxiliem nessa tarefa, portanto é necessário avaliar a disponibilidade de ferramentas de apoio à gestão das coleções digitais, à segurança, às políticas de gestão de direitos, entre outros.

Gestão/administração do repositório

- Gestão de coleções – o *software* possibilita a implementação de modelos diferenciados de gestão de coleção para o mesmo repositório (por exemplo, gestão por área de conhecimento, por tópicos, por comunidades, etc.)
- Estatísticas – verifica se o *software* possui módulo para a geração de estatísticas; identifica quais são os relatórios e gráficos oferecidos pelo sistema e em que padrões e formatos.
- Controle de direitos – verifica se o *software* possui módulo de apoio a gestão de direitos.
- Interface administrativa – avalia as interfaces para apoio a gestão segundo os parâmetros de ergonomia, usabilidade, acessibilidade, etc.
- Preservação digital – verifica se o software dispõe de dispositivos de apoio à gestão dos processos de preservação digital. Por exemplo, suporte à migração (ver também o item 3.5.2).

Gestão do usuário final

Circunscreve as facilidades e grau de autonomia que o *software* oferece aos usuários finais, particularmente à comunidade institucional, de criar subcomunidades específicas capazes de gerar seus próprios fluxos de trabalho e implementar políticas e mecanismos específicos de indexação, de recuperação e de exibição exclusiva de suas coleções.

- Ferramentas de gestão do usuário – avalia a disponibilidade de instrumentos de gestão voltados para o usuário final.

Segurança da informação

Compreende a capacidade do sistema de atender aos requisitos de segurança da instituição e assegurar a integridade física dos estoques de conteúdos digitais e intervenções e acessos indevidos.

- Autenticação e direitos de acesso – avalia se o *software* oferece funcionalidades para a gestão da segurança da informação: controle de acesso, autenticação de usuários, níveis de permissão, implementação de papeis para a gerencia de conteúdo e administração do sistema, etc.; uso de padrões: LDAP, X509, SSL.
- *Backup* e restauração – avalia disponibilidade de suporte aos processos de *backup* e restauração.
- Criptografia – avalia a disponibilidade de suporte a cifragem de dados durante a transmissão de conteúdos.

SOFTWARE LIVRES PARA REPOSITÓRIOS INSTITUCIONAIS: ALGUNS EXEMPLOS

Nesta seção discutiremos rapidamente as características técnicas de alguns dos muitos pacotes de *software* que possibilitem o desenvolvimento de repositórios institucionais. Nossa atenção estará voltada para os programas mais usados e/ou conhecidos no Brasil que estão

disponíveis sob licenças de fonte aberta, significando que eles estão disponíveis gratuitamente e podem ser livremente modificados, atualizados e redistribuídos.

É importante salientar que as características e versões dos programas, assim como os padrões utilizados, pelo próprio dinamismo da área, mudam frequentemente. As informações aqui registradas refletem as condições dos programas no momento em que o presente texto foi escrito[2].

DSPACE

URL: <http://www.dspace.org>

Descrição:
O DSpace é um projeto cooperativo de desenvolvimento liderado pelas bibliotecas do Massachusetts Institute of Technology (MIT) e pelos laboratórios da corporação Hewlett-Packard (HP), conduzido sob as diretrizes da DSpace Federation[3]. Na página do Dspace, o projeto é descrito como "um sistema de repositório digital inovador que captura, armazena, indexa, preserva e redistribui materiais de pesquisa em formato digital produzida por comunidades acadêmicas dentro do contexto de organizações de pesquisa e de universidades".

Disponibilidade:
O Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT)[4] disponibiliza para *downloading* gratuito em seu *site* uma versão em português do programa.

- Distribuição – *software* de fonte aberta distribuído sob a licença BSD open source.

---

2 Além das páginas *web* dos programas foram consultados trabalhos dos seguintes autores autores: Raym Crow (2004), Mary R. Barton (2004), Richard Wyles (2006).

3 <http://dspace.org/federation/index.html>

4 <http://www.ibict.br>

- *Downloading* – no site: <http://sourceforge.net/projects/dspace>; para a versão em português customizada pelo IBICT: <http:// dspace.ibict.br/ >

Características Técnicas

- Ambiente Operacional – Unix, Linux, Windows
- Tecnologias usadas – Java, Tomcat Servlet Engine
- Banco de Dados – PostgreSQL, MySQL, Oracle
- Motor de Pesquisa – Lucene ou Google
- Formatos aceitos – sem restrições
- Extensível via Java API

Padrões

- Interoperabilidade – Protocolo OAI-PMH, Web Services, SRU/SRW
- Esquema de metadados aceitos – Dublin Core qualificado
- Identificadores – Handle System
- Preservação digital – aderente ao modelo OAIS – Open Archive Information System; o *software* é focado no problema de preservação digital de longo prazo de materiais de pesquisa depositados.
- Importação/exportação de dados – formato XML e padrão METS.

Características específicas

- Implementa o conceito de comunidades
- Voltado para repositórios institucionais
- Foco em materiais para pesquisa e ensino.
- Workflow para submissão de conteúdos
- Interface *web* customizável

EPRINTS

URL: <http://www.eprints.org/>

Descrição:

O *software* Eprints tem a maior e a mais distribuída base instalada dentre os sistemas de repositórios digitais em operação; estima-se que ele seja o *software* que roda sob a maioria dos repositórios acadêmicos em todo mundo (MILLINGTON; NIXON, 2007). O EPrints foi desenvolvido na Universidade de Southampton, Inglaterra, por Rob Tansley e Christopher Guterridge membros da equipe de Stevan Harnard, pioneiro do movimento de acesso livre. A primeira versão do sistema foi publicamente lançada nos fins do ano 2000 como um *software* de fonte aberta, distribuído sob a licença GPL – GNU Public Licence.

Disponibilidade

O IBICT traduziu para o português e customizou o programa Eprints e o disponibiliza no seu *site* para downloding gratuito com o nome de Dialogo Científico (DICI).

- Distribuição – *software* de fonte aberta distribuído sob a licença GPL – GNU Public Licence;
- *Downloading* –

Características Técnicas

- Ambiente operacional – Unix/Linux
- Tecnologias usadas – Perl
- Banco de dados – MySQL
- Motor de pesquisa/indexação – MG/Managing Giga
- Extensível – via API usando a linguagem de programação PERL

Padrões Atendidos

- Interoperabilidade – Protocolo OAI-PMH;
- Esquema de metadados aceitos – Dublin Core;
- Identificadores – não implementa;
- Formatos aceitos – sem restrições
- Importação/exportação de dados – exportação de buscas: Dublin Core, METS, Reference manager, OpenURL Context Object e outros

Características específicas

- Alerta e exportação de resultado de buscas em formato RSS
- *Workflow*
- Alerta contra duplicatas

GREENSTONE

URL: <http://www.greenstone.org/>

Descrição:

O Greenstone – conforme está declarado em sua página oficial – é um conjunto de *software* projetado para criar e distribuir coleções digitais, proporcionando uma nova forma de organizar e publicar informações na Internet ou em CD-ROM (ou em outras mídias) na forma de biblioteca digital totalmente pesquisável e dirigida por metadados.

O *software* – que tem uma forte conotação social e humanitária, e pode ser considerado uma continuação da trajetória do Micro-ISIS – foi desenvolvido no âmbito do projeto New Zeland Digital Library[5] sediado na Universidade de Wikato; a sua distribuição é um trabalho cooperativo entre a UNESCO e a ONG Humanitarian Information for All ou Human Info[6], de origem Belga, que tem como objetivo prover informação para todas as pessoas envolvidas em desenvolvimento, bem estar social e necessidades básicas. O projeto se desenrolou no contexto das ações do Programa Informação para Todos da UNESCO, e tem como meta oferecer aos usuários, especialmente os vinculados às universidades, bibliotecas e outras instituições de conhecimento voltadas para o serviço público, uma ferramenta que os capacite a construir suas próprias bibliotecas digitais.

Disponibilidade:

- Distribuição – *software* de fonte aberta, distribuído sob a licença

5 <http://nzdl.sadl.uleth.ca/cgi-bin/library>

6 <http://humaninfo.org/home_flash.html>

GNU General Public License;
- *Downloading* – no *site* <http://www.greenstone.org/download>.

Características Técnicas
- Ambiente Operacional – Unix, Linux, Windows, Sun Solaris, Mac OS/X;
- Tecnologias usadas – Perl,Java,C++;
- Banco de Dados – GDBM (GNU Database Manager);
- Motor de Pesquisa/Indexação – MG; MG+/MGPP; Lucene.

Padrões atendidos
- Interoperabilidade – Protocolo OAI-PMH; Z39.50; Corba;
- Esquema de metadados aceitos – Dublin Core qualificado e não qualificado; RFC 1807; NZGLS;AGLS. Permite o uso de *plug-ins* para o uso de outros esquemas;
- Identificadores – não implementa;
- Formatos aceitos – sem restrições, realizado por meio de uso de *plug-ins* específicos;
- Importação/exportação de dados – padrão METS; formato CDS/ISIS, Dspace.

Características específicas:
- Fácil instalação;
- Multilíngue – quatro idiomas principais – inglês, francês, espanhol e russo - e adicionalmente mais de 25 interfaces de idiomas disponíveis, incluindo o português;
- Permite a geração de coleções digitais em CD-ROM;
- Editor de metadados;
- Exemplo de coleção – disponibiliza coleções de demonstração no site do programa.

FEDORA

URL: <http://www.fedora-commons.org/>

Descrição:

Fedora – sigla para Flexible Extensible Digital Object and Digital Repository Architecture – é um *software* de código aberto desenvolvido em conjunto por equipes da Cornell University e da University of Virginia, ambas nos Estados Unidos, e distribuído de acordo com os termos da Licença Pública Mozilla. O Fedora (que não deve ser confundido com a distribuição da Red Hat do sistema operacional Linux que tem o mesmo nome) teve o seu desenvolvimento iniciado em 1997 como um projeto de pesquisa liderado por Carl Lagoze e Sandy Payette, dois pesquisadores notáveis na área de bibliotecas digitais, com fundos da Defense Advanced Research Project Agency (DARPA) e da National Science Foundation (NSF), ambas organizações governamentais norte-americanas. Os resultados deste trabalho de pesquisa viabilizaram o desenvolvimento de uma arquitetura para construção e gestão de repositórios de objetos digitais complexos, concretizado pelo *software* Fedora. O desenvolvimento do Fedora contou também com o suporte financeiro da Fundação Andrew W. Mellon Foundation[7] (PAYETTE; STAPLES, 2002).

A arquitetura Fedora cria uma infraestrutura ampla para o armazenamento, gestão e disseminação de objetos digitais complexos, incluindo o relacionamento entre eles. O conceito central do sistema é um poderoso modelo de objeto digital que estabelece como unidade de informação o "objeto digital Fedora". O modelo apóia visões múltiplas de cada objeto e dos seus relacionamentos tornando o sistema uma ferramenta bastante flexível.

Disponibilidade:

- Distribuição – *software* livre, de fonte aberta, distribuída sob Licença Pública Mozilla;
- *Downloading* – no *site* <http://www.fedora-commons.org/software>.

Características Técnicas

- Ambiente Operacional – Sun Solaris, Unix, Linux e Windows
- Tecnologias usadas – Java
- Banco de Dados – McKoi, MySQL, Oracle 9 e Postgre SQL

7 <http://www.mellon.org/>

Padrões atendidos

- Interoperabilidade – OAI-PMH, Web Services, RDF
- Esquema de metadados aceitos – Dublin Core, Fedora FOXML
- Importação/exportação de dados – padrão METS, MPEG-21 DIDL, FOXML
- Identificadores persistentes – PID
- Preservação digital – internamente todos os objetos digitais Fedora são representados como arquivos XM, incluindo dados e metadados e relacionamentos. O *software* é aderente ao padrão OAIS/ISO.
- Formatos aceitos – sem restrições

Características específicas:

- Gerência de versão – o *software* armazena a história de todas as modificações sofridas pelo objeto digital
- Objeto digital – implementa o conceito de objeto digital complexo e o relacionamento entre eles

NOU-RAU

URL: <http://www.softwarelivre.unicamp.br/index.php/P%C3%A1gina_principal>

Descrição:

O sistema Nou-Rau, desenvolvido pelo Centro de Computação da Unicamp, tem como objetivo implementar um sistema *on-line* para arquivamento, indexação, acesso controlado e mecanismos eficientes para busca de documentos digitais. Para isso, o sistema recebe documentos digitais em diversos formatos, em seguida converte-os para texto puro, indexando-os através do uso da ferramenta de *software* ht://Dig – WWW Search Engine Software – disponível livremente em <http://www.htdig.org//>. O ambiente operacional nativo é o Linux ou outro da família UNIX. O Nou-Rau é um sistema de código aberto

distribuído sob a licença GPL <http://www.gnu.org/copyleft/gpl.html> e pode ser baixado livremente a partir da página *web* do programa.

A informação no ambiente Nou-Rau é organizada em áreas denominadas “tópicos”, que representam um assunto específico e serve para agrupar documentos relacionados. Os tópicos são organizados hierarquicamente possibilitando o estabelecimento de subtópicos. Por exemplo, o tópico “Energia Nuclear” e o subtópico “Medicina Nuclear”. É possível nomear um gestor para cada tópico que será responsável, entre outras coisas, pela aprovação dos documentos submetidos.

Disponibilidade:

- Distribuição – um sistema de código aberto distribuído sob a licença GPL- GNU Public Licence;
- *Downloading* – no *site* <http://www.softwarelivre.unicamp.br/index.php/P%C3%A1gina_principal>.

Características Técnicas

- Ambiente Operacional – Linux
- Tecnologias usadas – linguagem PHP e Perl, servidor web Apache
- Banco de Dados – Postgre SQL
- Ferramenta de busca – busca em texto completo via ht://Dig;

Padrões atendidos

- Interoperabilidade – OAI-PMH, captura de dados via Z39.50
- Formatos aceitos – o sistema recebe documentos digitais em diversos formatos, em seguida converte-os para texto puro via um conjunto de programas livres.

Características específicas

- Indexação – o sistema recebe documentos digitais em diversos formatos, em seguida converte-os para texto puro, indexando-os através do uso da ferramenta de *software* ht://Dig; o sistema é capaz de indexar conteúdos de documentos nos formatos ASCII, HTML, RTF, SGML XML, MS Word, Power Point, PDF, Post Script, TeX, Látex e DVI.

- Organização da informação em tópicos que podem ser geridos de forma autônoma.
- Implementa ainda um conjunto pré-definido de papéis que os diversos tipos de usuários podem desempenhar no domínio do sistema

Exemplos

- Biblioteca digital da Unicamp: <http://libdigi.unicamp.br/>

À GUISA DE CONCLUSÃO

Os repositórios digitais surgem como uma expressão real das mudanças impulsionadas pela apropriação por parte da comunidade acadêmica de conhecimentos, tecnologias e padrões das áreas de TI para constituição de novos patamares de publicação, comunicação e cooperação. Uma evidência eloquente desse fenômeno é o protocolo Open Archive voltado para interoperabilidade de repositórios digitais acadêmicos que tem origem no próprio seio da comunidade acadêmica.

As soluções de *software* para a criação de repositórios digitais também seguem esse principio. Surpreendentemente os programas tecnicamente mais sofisticados foram desenvolvidos nas universidades e em centros de pesquisa, resultados de projetos de pesquisa, da colaboração entre a universidade e empresas ou movidos para a solução de problemas. Como desdobramento, esses programas são distribuídos livremente e têm os seus códigos abertos, permitindo dinâmicas interessantes de desenvolvimento e troca de experiências.

A ampla oferta de *software* livres de qualidade, baseados em padrões abertos e apoiadas por comunidades de desenvolvedores criaram condições ideais para a implementação, a um custo relativamente baixo, de um grande número de repositórios digitais de toda natureza – bibliotecas, arquivos, museus digitais, repositórios temáticos e institucionais e outros.

Entretanto, essa oferta ampla de *software* livres para desenvolvimento de repositórios digitais impõe às comunidades interessadas o

desafio de selecionar qual deles se enquadra melhor nas necessidades dos usuários, na infraestrutura disponível, na expertise para gerenciar e operar o repositório e no orçamento da instituição. A seleção se torna crítica na medida em que cada um das plataformas de *software* disponíveis foi desenvolvido em contextos distintos e têm objetivos e vocações específicas.

O presente texto limitou-se a alinhar alguns critérios técnicos e funcionais que podem subsidiar o gestor na sua decisão, entretanto, fica a necessidade de desenvolver modelos e metodologias mais formais, de amplo espectro de aplicação para avaliações de toda a infraestrutura tecnológica necessária à implementação do que parece ser um fenômeno importante para a área acadêmica, que são os repositórios institucionais.

REFERÊNCIAS

BARTON, Mary R.; WATERS, Margareth M. **Creating an institutional repository**: LEADIRS Workbook. MIT Libraries, 2004. Disponível em: <http://www.dspace.org/images/stories/leadirs.pdf>. Acesso em: 10 jan. 2010.

CÉSPEDES, Zulia Ramirez. Criterios e indicadores para evaluar las bibliotecas digitales. 2006. **Acimed: revista cubana de los profesionales de la información y la comunicación en salud**, v. 14, n. 6, 2006. Disponível em: <http://bvs.sld.cu/revistas/aci/vol14_6_06/aci04606.htm>. Acesso em: 25 jan. 2010.

CROW, Raym. **A guide to institutional repository software**. 3 ed. New York: Open Society Institute, 2004. Disponível em: <http://www.soros.org/openaccess/pdf/OSI_Guide_to_IR_Software_v3.pdf>. Acesso em 30 dez. 2009.

LINCH, Clifford A. Institutional repositories: essential infrastructure for scholarship in the digital age. **ARL**, n. 226, p. 1-7, Feb. 2003. Disponível em: <http://scholarship.utm.edu/21/1/Lynch,_IRs.pdf>. Acesso em: 10 jan. 2010.

MARILL, Jennifer L.; LUCZAK. Evaluation of digital repository software at the National Library of Medicine. **D-Lib Magazine**, v. 15, n. 5/6, May/June 2009. Diponível em: <http://www.dlib.org/dlib/may09/marill/05marill.html>. Acesso em: 30 dez. 2009

MILLINGTON, Peter; NIXON, Willian J. Eprints 3 Pre-Launch Briefing. **Ariadne**, v. 50, Jan. 2007. Disponível em <http://www.ariadne.ac.uk/issue50/eprints-v3-rpt/>. Acesso em: 10 jan. 2010.

PAYETTE, Sandra; STAPLES, Thornton. The Mellon Fedora Project: Digital

Library Architecture Meets XML and Web Services, In: **Research and Advanced Technology for Digital Libraries**. Berlin / Heidelberg : Springer, September 2002. p. 87-106.

POWELL, Andy. **Notes about possible technical criteria for evaluating institutional repository (IR) software**. UKOLN, Dec. 2005. Disponível em: <http://www.ukoln.ac.uk/distributed-systems/jisc-ie/arch/ir-software.pdf>. Acesso em: 09 jan. 2010.

SAYÃO, Luis Fernando. Interoperabilidade das bibliotecas digitais: o papel dos sistemas de identificadores persistentes - URN, PURL, DOI, Handle System, CrossRef e OpenURL. Transinformação, v. 19, n. 1, 2007. Disponível em: <http://revistas.puc-campinas.edu.br/transinfo/include/getdoc.php?id=469&article=245&mode=pdf >. Acesso em: 10 jan. 2010.

_______. Preservação digital no contexto das bibliotecas digitais. In:

MARCONDES, Carlos Henrique; KURAMOTO, Helio; TOUTAIN, Lidia Brandão; SAYÃO, Luis Fernando (Org.). **Bibliotecas digitais**: saberes e práticas. Salvador; Brasília: UFBA; IBICT, 2006. p. 115-149.

SAYÃO, Luis Fernando; MARCONDES, Carlos Henrique. O desafio da interoperabilidade e as novas perspectivas para as bibliotecas digitais. **Transinformação**, v. 20, n. 2, 2008. Disponível em: <http://revistas.puc-campinas.edu.br/transinfo/viewarticle.php?id=215>. Acesso em: 10 jan. 2010.

WYLES, Richard. **Technical evaluation of open source repository solutions**. Christchurch Polytechnic Institute of Technology (CPIT), 2006. Disponível em: <https://eduforge.org/docman/view.php/131/1062/Repository%20Evaluation%20Document.pdf>. Acesso em: 30 dez. 2009.

Arquitetura da informação para repositórios científicos digitais

Liriane Soares de Araújo de Camargo
Silvana Aparecida Borsetti Gregorio Vidotti

INTRODUÇÃO

Ambientes informacionais digitais estão surgindo com o intuito de possibilitar o gerenciamento, o tratamento, a recuperação, o uso, a preservação e a disseminação de informações e de documentos científicos e acadêmicos. Nesse contexto, os repositórios digitais têm sido reconhecidos como um tipo desses ambientes, visando a contribuir de forma significativa na comunicação entre as comunidades científicas.

Os repositórios digitais podem ser considerados como locais de armazenamento de coleções digitais de uma determinada instituição ou comunidade e utilizam sistemas de informação que possibilitam funções como: criação de comunidades e de coleções, cadastro de usuários, gerenciamento de políticas de conteúdos e auto-arquivamento de documentos.

No âmbito das instituições de ensino superior e de pesquisa, os repositórios digitais possibilitam o controle, a preservação e a visibilidade da produção científica, minimizando custos de publicações e permitindo o acesso irrestrito de outras comunidades universitárias e de pesquisa e da sociedade em geral.

O desenvolvimento de repositórios digitais pode ser auxiliado por uma Arquitetura da Informação (AI) que considere a complexidade e a grande quantidade de itens envolvidos no processo de desenvolvimento de tais ambientes, bem como no processo de representação e de recuperação de itens documentários.

A Arquitetura da Informação oferece diretrizes e informações necessárias para auxiliar o desenvolvimento de ambientes informacionais, abordando processos de estruturação, organização, representação, recuperação, navegação, apresentação e disseminação de conteúdos e serviços.

Baseado nesse contexto apresenta-se uma Arquitetura da Informação para Repositórios Digitais a fim de facilitar o desenvolvimento desse tipo de ambiente científico e aumentar a usabilidade do ambiente e das informações nele contidas.

A Arquitetura da Informação proposta considera processos de acessibilidade, usabilidade, qualidade de *software* e personalização para a melhoria da qualidade de acesso e de uso do próprio ambiente e dos trabalhos científicos e acadêmicos. Vale ressaltar que "processos" são considerados neste texto como diretrizes ou princípios que podem ser utilizados por desenvolvedores e arquitetos da informação na construção de ambientes digitais.

A partir do contexto exposto, pode-se afirmar que ambientes científicos como os repositórios digitais podem ser mais explorados no que se refere ao aperfeiçoamento de recursos interativos e colaborativos, principalmente os de personalização e de customização, a fim de recuperar e disseminar informações científicas para comunidades de pesquisa e para a sociedade em geral.

ARQUITETURA DA INFORMAÇÃO

A Arquitetura da Informação é uma área ainda em consolidação, em que se podem identificar pontos em comum e divergentes entre os autores. Entre os pontos em comum encontram-se o oferecimento de informações para auxiliar o arquiteto da informação no desenvolvimento de ambientes digitais e na definição de Richard Saul Wurman, que popularizou o termo 'Arquitetura da Informação' em meados da década de 60, com a qual vários autores da área concordam ser: "uma estrutura ou mapa de informação, permitindo que as pessoas/usuários encontrem seus caminhos pessoais para o conhecimento" (WURMAN,

1996, 2001). Já entre os pontos divergentes, distintos autores apresentam essas informações como diretrizes, princípios, planos, sistemas, guias, regras, critérios etc., não havendo um consenso na literatura em relação a esse aspecto.

Nesse sentido, Santos (2001, p. 3) relata que a Arquitetura de Informação deve ser encarada como "uma das maneiras de se aprimorar a usabilidade de um sistema através do desenvolvimento de uma estrutura de informação que possibilite ao usuário atingir suas metas de interação durante o processo de busca de informação, seja em um website de compras, de entretenimento, de informação, comunitário ou outros".

Wurman (1996) considera ainda que os arquitetos da informação devem entender os problemas de reunião, organização e apresentação da informação como análogos aos de um arquiteto ao projetar um edifício que atenda às necessidades de seus ocupantes. De forma complementar a essa afirmação, Siqueira (2008, p. 30) relata que "a visão de Wurman é derivada de sua formação como arquiteto e seu principal propósito é estender os conceitos chaves de organização de espaços, desenvolvidos na arquitetura, para os espaços informacionais".

McGee e Prusak (1994, p. 129) afirmam que o objetivo de uma Arquitetura da Informação é "criar um 'mapa' abrangente dos dados organizacionais e em seguida construir um sistema baseado nesse mapa". Assim, a AI deve ser a planta e o mapa para a organização virtual da informação, e deve incluir as formas como o usuário navegará e acessará a informação. Esses autores (1994, p.129) relatam ainda que o termo Arquitetura da Informação "é um termo complexo, pois combina duas palavras que possuem uma vasta gama de conotações". Entretanto, o importante é compreender que a AI aborda a construção de ambientes digitais. Nesse sentido, apresentam-se a seguir algumas definições:

- Latham (2002, p. 825, tradução nossa) relata que "o termo 'arquitetura da informação', como é atualmente utilizado, é tipicamente aplicado ao projeto e desenvolvimento de websites". O autor explica que o termo, contudo, pode ser aplicado mais amplamente para projeto e desenvolvimento de sistemas e pro-

dutos informacionais de forma geral, desde que envolva o usuário e a coordenação de numerosos componentes técnicos, incluindo banco de dados, metadados, gerenciamento de conteúdos dinâmicos, múltiplas mídias, fonte única e modelos de informação.

- Sotillos (2002, p. 35) sustenta que a AI não trata somente de projeto de sites (os espaços de informação) na WWW e sim, em geral, de qualquer ambiente digital.
- Para Vidotti, Cusin e Corradi (2008, p. 182), a

> Arquitetura da Informação enfoca a organização de conteúdos informacionais e as formas de armazenamento e preservação (sistemas de organização), representação, descrição e classificação (sistema de rotulagem, metadados, tesauro e vocabulário controlado), recuperação (sistema de busca), objetivando a criação de um sistema de interação (sistema de navegação) no qual o usuário deve interagir facilmente (usabilidade) com autonomia no acesso e uso do conteúdo (acessibilidade) no ambiente hipermídia informacional digital.

- Camargo (2010, p. 48) define Arquitetura da Informação como

> Uma área do conhecimento que oferece uma base teórica para tratar aspectos informacionais, estruturais, navegacionais, funcionais e visuais de ambientes informacionais digitais, por meio de um conjunto de procedimentos metodológicos a fim de auxiliar no desenvolvimento e no aumento da usabilidade de tais ambientes e de seus conteúdos.

Vale ressaltar que concorda-se com Siqueira (2008, p. 33) quando ele afirma que "não é possível delimitar a Arquitetura da Informação ao uso pragmático de tratamento de documentos, muito menos, restringi-la ao contexto da criação de sítios na Internet". "A aplicação da Arquitetura da Informação viabiliza a redução do custo de acesso à informação, potencializando o seu valor para o usuário".

Assim, os ambientes informacionais digitais devem ser desenvol-

vidos a fim de atender tanto as necessidades da instituição quanto as necessidades dos usuários, e o uso da Arquitetura da Informação no projeto e desenvolvimento desses ambientes pode aumentar e melhorar a sua usabilidade e permitir a acessibilidade.

REPOSITÓRIO CIENTÍFICO DIGITAL

Os repositórios científicos digitais, caracterizados como um tipo de ambiente informacional digital, possibilitam a interoperabilidade de dados, o controle e o armazenamento da produção científica, a preservação da informação a longo prazo, o auto-arquivamento do documento, o acesso livre, a recuperação e a disseminação da informação científica, dando visibilidade à produção científica e minimizando os custos de publicação.

> Para Viana, Márdero Arellano e Shintaku (2006, p. 3)
> Um repositório digital é uma forma de armazenamento de objetos digitais que tem a capacidade de manter e gerenciar material por longos períodos de tempo e prover o acesso apropriado. Essa estratégia foi possibilitada pela queda nos preços no armazenamento, pelo uso de padrões como o protocolo de coleta de metadados da Iniciativa dos Arquivos Abertos (OAI-PMH), e pelos avanços no desenvolvimento dos padrões de metadados que dão suporte ao modelo de comunicação dos arquivos abertos.

Rodrigues (2005) ressalta que os repositórios podem ser de dois tipos: disciplinar ou institucional. Os disciplinares são sistemas abertos que arquivam os resultados de investigação de uma ou várias disciplinas e também são conhecidos como repositórios temáticos. O repositório institucional é a reunião de todos os repositórios temáticos hospedados em uma organização.

Café e outros (2003) explicam que no caso de uma universidade, cada departamento trata de uma área do conhecimento e, portanto,

seu repositório temático será específico no assunto deste departamento. A união de todos os repositórios das diversas unidades de pesquisa comporá o repositório institucional, caracterizando-o como multidisciplinar. Os autores afirmam ainda que

> Um repositório temático se constitui em um conjunto de trabalhos de pesquisa de uma determinada área do conhecimento, disponibilizados na Internet. Esses repositórios utilizam tecnologias abertas e seguem a filosofia da Iniciativa dos Arquivos Abertos, promovendo a maior acessibilidade à produção dos pesquisadores e à discussão entre seus pares. Suas principais características são: processamento automático dos mecanismos de discussão entre os pares; geração de versões de um mesmo documento; tipologia variada de documentos; auto-arquivamento; e interoperabilidade entre todos os repositórios temáticos e seus serviços agregados. (CAFÉ et al., 2003, p. 3)

Segundo Crow (2002) os repositórios institucionais podem ser coleções digitais que armazenam, preservam, divulgam e dão acesso à produção intelectual de comunidades universitárias. O autor relata que os repositórios pretendem intervir e dar resposta a duas questões estratégicas que as universidades enfrentam: contribuir para o aumento da visibilidade, do estatuto, da imagem e do "valor" público da instituição, servindo como indicador tangível da qualidade dessa universidade e demonstrando a relevância científica; e contribuir para a reforma do sistema de comunicação científica, expandindo o acesso aos resultados da investigação, reassumindo o controle acadêmico sobre a publicação científica, aumentando a competição e reduzindo o monopólio das revistas científicas – que se traduziria também em economias para as universidades e as bibliotecas que as servem.

Em 2005 foram debatidas definições de repositórios pelos membros da Coalition for Networked Information (CNI) que envolveram duas visões com ênfases diferentes: a primeira caracteriza o repositório como primariamente uma maneira de disseminação de várias for-

mas de e-print[1] para trabalhos universitários, a segunda abordagem o conceitua como um lugar amplo de documentação de trabalhos intelectuais (pesquisa e ensino).

Atualmente, pode-se considerar que o repositório digital surgiu com propósitos de preservação da memória e visibilidade institucional, porém, esses ambientes não precisam ser obrigatoriamente científicos. Eles se iniciaram com esse objetivo, mas podem ser desenvolvidos com fins administrativos (visando a comunidade funcional). O repositório digital é um ambiente recente que deverá sofrer mudanças conceituais ao decorrer do tempo, no entanto deve manter em seu princípio a preservação da memória a longo prazo.

Os repositórios digitais podem ser definidos como "coleções digitais que armazenam, preservam e tornam disponível a produção intelectual de uma ou mais universidades, sem qualquer custo para o produtor e consumidor da informação" (SARMENTO et al., 2005, p. 3).

Para o desenvolvimento de repositórios há a necessidade de definições de políticas institucionais, considerando a elaboração de acervos originalmente digitais A política influencia o processo de auto-arquivamento, que necessariamente não significa auto-publicação. Esse processo, geralmente, consiste na realização de *login*, na escolha da comunidade e da coleção e nas etapas de descrever, carregar, verificar, licenciar e completar. Os repositórios têm processos e funções similares às bibliotecas digitais, entretanto, eles possibilitam o auto-arquivamento e a interoperabilidade entre diversos sistemas de informação por meio da coleta de metadados em arquivos abertos.

Apresentam-se alguns exemplos de repositórios científicos digitais, em especial os institucionais: o ArXiv[2] – repositório institucional interdisciplinar de *e-prints;* o Australian Research Repositories Online to the World (ARROW)[3] – da Biblioteca Nacional da Austrália; o The Scholarly Publishing and Academic Resources Coalition (SPARC)[4] – que

1 Versão digital de um documento de pesquisa

2 http://arxiv.org/

3 http://www.arrow.edu.au/

4 http://www.arl.org/sparc/repositories/

apresenta coleções digitais baseadas em universidades; o E-prints in Library and Information Science (E-LIS)[5] – um repositório temático que reúne a produção das áreas de Biblioteconomia, de Ciência da Informação e áreas afins; e o Repositório da Universidade do Minho (RepositoriUm)[6] –cujo propósito é armazenar, preservar e divulgar a produção intelectual das diversas comunidades científicas da Universidade do Minho – Portugal.

Os repositórios digitais utilizam ferramentas de auxílio ao seu desenvolvimento como o EPrints – Digital Repository Software[7] e o DSpace[8]. A primeira ferramenta citada, conhecida como Eprints, fornece uma base para as instituições acadêmicas criarem seus próprios repositórios de pesquisa. O DSpace advém de um resultado de esforço conjunto entre o Massachusetts Institute of Technology (MIT) e a Hewlett-Packard (HP) e consiste em um *software* que pode ser modificado, melhorado e distribuído livremente, destinado a receber, preservar e redistribuir a produção intelectual de investigação de universidades e outras instituições em formatos digitais.

Além dessas ferramentas podem-se encontrar outras como: Fedora,[9] da University of Virginia e Cornell University; i-Tor,[10] do Netherlands Institute for Scientific Information Service; Archimede,[11] da Laval University Library; e a CDSware,[12] da European Organization for \ Nuclear Research (CERN).

AMBIENTES CIENTÍFICOS DIGITAIS

Para desenvolver uma Arquitetura da Informação para Repositórios Científicos Digitais foi necessária a realização de uma análise em vários

5 http://eprints.rclis.org

6 http://www.uminho.pt/default.aspx

7 http://www.eprints.org/software/

8 http://www.dspace.org

9 http://www.fedora.info/

10 http://www.i-tor.org/en/toon

11 http://archimede.bibl.ulaval.ca/archimede/pages/home/index.jsf

12 http://cdsware.cern.ch

tipos de ambientes científicos digitais a fim de identificar características e recursos específicos. Esses ambientes abrangem: bibliotecas digitais, periódicos científicos e repositórios científicos digitais. A identificação dessas características e recursos visa à elaboração de uma listagem de itens a serem analisados no contexto da Arquitetura da Informação, com o intuito de auxiliar desenvolvedores e usuários na avaliação de ambientes informacionais digitais, principalmente, dos repositórios científicos digitais.

Pode-se considerar que existem objetivos em comum entre esses tipos de ambientes informacionais: armazenar, representar, preservar, interoperar, facilitar o acesso e disseminar informações. Assim, várias atividades desses ambientes se assemelham: oferecer acesso remoto e simultâneo, gerenciar conteúdo, preservar e recuperar as informações, tratar o conteúdo considerando a representatividade, a segurança e a confiabilidade do mesmo, oferecer coleções de documentos bibliográficos e completos, oferecer produtos e serviços, utilizar metadados e possuir diversas fontes e formatos.

Pode-se considerar ainda que os tipos de ambientes selecionados para análise contemplam de forma expressiva as características gerais de todos os tipos de ambientes científicos digitais, entretanto, esses ambientes podem incorporar novas funções, serviços e princípios para atender da melhor forma possível os seus usuários, considerando a dinamicidade e a interatividade da Internet. Por exemplo: a recuperação da informação pode ser auxiliada por um vocabulário controlado e pelo tratamento semântico de acordo com a comunidade efetiva, e os metadados podem ser adaptados para descrever objetos digitais de acordo com o acervo do ambiente, bem como contribuir para a utilização de um agregador de conteúdo.

Apresentam-se a seguir alguns elementos essenciais que devem ser abordados pelos ambientes científicos digitais, principalmente pelos repositórios.

- **Ferramenta de Busca**: também conhecidas como pesquisadores, mecanismos ou motores de busca, são programas computacionais desenvolvidos com o objetivo de registrar, em bases de dados, as representações descritivas e temáticas dos websites, com a finalidade de possibilitar a recuperação de informações

solicitadas, segundo as estratégicas de busca adotadas pelos usuários. A estratégia de busca depende do tipo de usuário e da própria ferramenta, pois esse mecanismo pode possibilitar uma estratégia simples e/ou avançada. Rosenfeld e Morville (1998) comentam sobre o sistema de busca, que demonstra a variedade de expectativas dos usuários, que podem: buscar por itens conhecidos, quando algumas necessidades são claramente definidas e requerem uma resposta simples, buscar por ideias abstratas (o usuário sabe o que ele quer, mas tem dificuldade em descrever), buscar de forma exploratória (o usuário sabe como expressar sua questão, mas não sabe exatamente o que espera encontrar e está apenas explorando uma questão para poder aprender algo mais), e buscar de forma compreensiva (os usuários querem todas as informações disponíveis sobre um determinado assunto).

- **Metadados**: segundo Alves (2010) os "metadados são atributos que representam uma entidade (objeto do mundo real) em um sistema de informação". A autora afirma que metadados "são elementos descritivos ou atributos referenciais codificados que representam características próprias ou atribuídas às entidades"; "são ainda dados que descrevem outros dados em um sistema de informação, com o intuito de identificar de forma única uma entidade (recurso informacional) para posterior recuperação". Os metadados trazem diversas vantagens para os usuários, pois por meio de uma representação padronizada dos recursos informacionais disponíveis em meio eletrônico, proporcionam o acesso mais amplo aos conteúdos, facilitam a busca, integram e compartilham recursos heterogêneos (GILLILAND-SWETLAND, 1999; ORTIZ-REPISO JIMÉNEZ, 1999). Para os profissionais da Ciência da Informação, o termo metadados está relacionado com o tratamento da informação, mais especificamente às formas de representação de um recurso informacional para fins de identificação, localização, preservação, administração e recuperação, ou seja, dados sobre catalogação e indexação que servem para organizar e tornar a informação mais acessível (GILLILAND-SWETLAND, 1999).

- **Política**: a maioria dos ambientes científicos digitais é desenvolvida por iniciativas de instituições responsáveis e confiáveis, as quais definem políticas que abrangem formas de uso e gerenciamento informacional e questões sobre direitos autorais. As políticas orientam as coleções e garantem a visibilidade do ambiente, prevendo: forma de acesso, tipo de documentos, restrições ao nível do conteúdo do documento, formas de depósito de documentos, tipologias de formatos, formato de documentos, digitalização de documentos, normalização de formatos, segurança e preservação da informação e normalização para documentos eletrônicos. Cada política varia de acordo com o tipo de ambiente e deve ser adequada ao contexto do objeto informacional. É importante ressaltar que os ambientes informacionais precisam adotar políticas e diretrizes de auto-avaliação e auto-reajuste/realimentação, bem como ter transparência na divulgação de resultados e utilizar critérios de reconhecimento pessoal para impulsionar a motivação para o sucesso do ambiente informacional.
- **Interoperabilidade**: normas, padrões e regras sempre foram desenvolvidos, a fim de garantir a precisão dos recursos informacionais para um acesso e recuperação efetivos, sobretudo em ambientes informacionais específicos, tais como as bibliotecas e repositórios digitais. A interoperabilidade é a capacidade de compartilhamento de informações em diferentes sistemas por meio de ferramentas como linguagem de marcação adequada como XML (Extensible Markup Language), uso de metadados e arquiteturas de metadados. As informações registradas e armazenadas em diferentes estruturas e comunidades do conhecimento poderão ser intercambiadas, possibilitando um trabalho conjunto entre sistemas e usuários.
- **Preservação**: muitos ambientes digitais não possuem URL (Uniform Resource Locator) permanente e de acordo com Coelho (2005, p. 9) algumas plataformas de desenvolvimento pode garantir "que a referência (URL) permaneça da mesma forma a longo prazo, pois os utilizadores necessitam de referências

permanentes e estáveis para os seus trabalhos e estes tornam-se fundamentais para as suas citações". Assim, a permanência das URLs é uma forma de preservação da informação, que segundo Boeres e Márdero Arellano (2005, p.2) "é a parte mais longa e também a última do ciclo de gerenciamento de objetos digitais, com ela é garantido o emprego de mecanismos que permitem o armazenamento em repositórios de objetos digitais e que garantem a autenticidade e perenidade dos seus conteúdos". Esses autores (2005, p. 4) relatam que preservação digital requer "estratégias e procedimentos para manter sua acessibilidade e autenticidade através do tempo, podendo requerer colaboração entre diferentes financiadoras e boa prática de licenciamento, metadados e documentação, antes de aplicar questões técnicas". Boeres e Márdero Arellano (2005, p. 10) esclarecem que

> Uma aceitável política de preservação digital implica em observar e aplicar procedimentos que podem ser inclusive aceitos como estratégias de preservação. Entre eles estão os relativos à tecnologia da informação, mais especificamente no tocante a compatibilidade de hardware, software e migração dos dados (conversão para outro formato físico ou digital, emulação tecnológica e espelhamento dos dados); observação da integridade do conteúdo intelectual a ser preservado; análise dos custos envolvidos no processo; o desenvolvimento de uma criteriosa política de seleção do que será preservado e, intimamente atrelado a isto, a observação das questões concernentes ao direito autoral.

Para o Commission On Preservation & Access/ Research Libraries Group as principais estratégias para a preservação digital estão relacionadas a preservação da tecnologia, a tecnologia de emulação e a migração da informação (WATERS; GARRETT, 1996). Além dessas estratégias, Sayão (2008, p. 176) cita a "preservação física, lógica, intelectual, do aparato e o monitoramente e a instrumentalização da comunidade-alvo". Em 2003 o grupo RLG-OCLC (Research Library Group) realizou uma pesquisa

com aspectos práticos da implementação de metadados de preservação e projetos de preservação digital. Esses metadados informam sobre a origem do material, os detalhes técnicos dos registros (como qual foi a versão do *software* usado, como foi construído o registro etc.) Isso pode ser usado como um meio de estocar a informação técnica que apóia a preservação dos objetos digitais e visa apoiar e facilitar a retenção, a longo prazo, da informação digital.

- **Acessibilidade**: iniciativas governamentais surgem para minimizar problemas de acesso e inclusão digital, visando auxiliar usuários portadores de necessidades especiais por meio de recomendações de princípios específicos de acessibilidade. Torres e outros (2002) relatam que a acessibilidade consiste em tornar disponível ao usuário, de forma autônoma, toda a informação que lhe for franqueável, independentemente de suas características corporais (individuais/orgânicas), sem prejuízos quanto ao conteúdo da informação. O autor relata ainda que devem ser feitas adequações de requisitos para usuários com limitações associadas à motricidade, audição e visão. Baranauskas e Mantoan (2001, p. 14) comentam que "aspectos de acessibilidade em páginas web consideram a variedade de contextos de interação que podem estar relacionados a diversos tipos de situações dos usuários com ou sem deficiência. Entre esses cidadãos encontra-se também a população de idosos". Freire e Fortes (2004) relatam que durante o desenvolvimento de um ambiente informacional digital é necessário que sejam levados em consideração os diferentes cenários em que o usuário poderá acessá-lo, tais como casos em que o usuário possui dificuldade para ler, ouvir, ou compreender o conteúdo do website, ou casos em que o usuário utiliza dispositivos com interfaces não convencionais. Também deve ser considerado que o usuário pode estar utilizando *browsers* e/ou sistemas operacionais diferentes, ou ter restrições quanto à velocidade da conexão com a Internet. Nesse sentido, Corradi e Vidotti (2007, p. 3) relatam que

> Para a efetivação da inclusão digital e social de usuários info-excluídos das potencialidades da era informacional, além de ampliar o acesso à Internet em todos os segmentos da sociedade, torna-se necessário promover ambientes informacionais que agreguem elementos de acessibilidade favoráveis à interação homem-computador.

As autoras (2007, p. 3) consideram que um "[...] ambiente digital pode possibilitar o atendimento às distintas formas de interação do usuário com a informação, respeitando suas condições sensoriais, lingüísticas e motoras em relação ao *hardware* e ao *software* utilizados". Winckler e Pimenta (2002, p. 2) relatam que "a maioria das recomendações ergonômicas e recomendação para acessibilidade não limita a utilização da interface apenas a pessoas com necessidades especiais". Algumas das recomendações podem ser úteis para qualquer usuário, como: descrever imagens e animações (atributo 'alt'), incluir transcrição de áudio e descrição de vídeos, usar cabeçalho, listas e estruturas consistentes etc. As recomendações de acessibilidade para a Web podem ser encontradas nos guias de acessibilidade do World Wide Web Consortium (W3C): Web Content Accessibility Guidelines (WCAG), Authoring Tool Accessibility Guidelines (ATAG) e User Agent Accessibility Guidelines (UAAG).[13]

- **Usabilidade**: esse termo começou a ser utilizado no início da década de 80, principalmente nas áreas de Psicologia Cognitiva e Ergonomia, como um substituto da expressão "*userfriendly*", considerando-o como facilidade de aprendizagem, rapidez no desempenho da tarefa, baixa taxa de erro, interface adequada ao sistema e satisfação subjetiva do usuário. Para Silvino e Abrahão (2003, p. 13) "a usabilidade, aferida pelos critérios ergonômicos e de funcionalidade, indica o grau de facilidade que a página oferece ao ser acessada". A usabilidade, por sua vez, refere-se à qualidade de interação entre os usuários e os ambientes informacionais digitais no momento do uso e está relacionada

13 http://www.w3.org/WAI/

> à Arquitetura da Informação, pois permite a avaliação desses ambientes em todas as fases de desenvolvimento, sob a ótica dos usuários e dos projetistas (VECHIATO, 2010). Os problemas de usabilidade mais recorrentes correspondem: a falta de atualização do ambiente, a interação usuário-sistema deficitária, a falha na navegabilidade e nas funcionalidades, a ausência de suporte e *feedback,* a dificuldade em acessar a informação desejada, as interfaces complicadas e de difícil uso e a ocorrência de erros. Vale comentar que existem muitas ferramentas disponíveis na web para avaliação de usabilidade e de desempenho de websites (por exemplo: ErgoList,[14] analisa recursos de usabilidade e de ergonomia por meio de questões, e Free Webmaster Tools,[15] que tem diversas ferramentas para confirmar se há *links* quebrados, sugerir palavras-chave etc). Contudo, alguns requisitos específicos de usabilidade podem ser subjetivos, sendo necessária uma análise específica e aprofundada que pode envolver usuários.

Além desses elementos, podem-se utilizar os princípios dos sistemas de organização, de busca, de navegação, de rotulagem e de representação apresentados na arquitetura da informação de Morville e Rosenfeld (2006).

ARQUITETURA DA INFORMAÇÃO PARA REPOSITÓRIOS CIENTÍFICOS DIGITAIS: PROCESSOS ESPECÍFICOS

Baseado nos princípios da Arquitetura da Informação e na análise realizada em ambientes científicos digitais, principalmente nos repositórios digitais, apresentam-se a seguir processos que podem ser aplicados para auxiliar os arquitetos da informação no desenvolvimento e na avaliação de repositórios digitais.

---

14 Disponível em: <http://www.labiutil.inf.ufsc.br/ergolist/)>

15 Disponível em: <http://www.iwebtool.com/tools/)>

PROCESSOS ESPECÍFICOS PARA REPOSITÓRIOS DIGITAIS

O quadro 1 apresenta alguns processos específicos para serem utilizados no desenvolvimento de ambientes informacionais como os repositórios digitais. Tais processos foram identificados a partir da observação direta não participativa em que foram observados serviços, recursos e funcionamento de vários repositórios digitais por meio de descrição de atividades, interações, ferramentas e outros itens relevantes, não havendo modificações nos mesmos.

<table>
<tr><th colspan="2">Processos específicos para Repositórios Digitais</th></tr>
<tr><td colspan="2">Possibilitar ao administrador a criação de comunidades</td></tr>
<tr><td colspan="2">Possibilitar ao administrador a criação de coleções</td></tr>
<tr><td colspan="2">Armazenar coleções digitais, suportando vários tipos e formatos de documentos</td></tr>
<tr><td colspan="2">Estabelecer utilizadores, suportando diversos tipos de atores</td></tr>
<tr><td colspan="2">Possibilitar edição de perfis, permitindo o cadastro e alterações de dados pessoais dos usuários</td></tr>
<tr><td rowspan="9">Estabelecer políticas</td><td>de acesso, que estabelece funções para os utilizadores</td></tr>
<tr><td>de tipo de documentos, que determina quais tipos de documentos serão armazenados</td></tr>
<tr><td>de restrições ao nível do conteúdo do documento, que estabelece restrições relativas ao conteúdo dos documentos pelos seus autores, nomeadamente a pesquisa, leitura, descarregamento (download), impressão e cópia</td></tr>
<tr><td>de depósito de documentos, que estabelece formas de depósitos</td></tr>
<tr><td>de tipologias de formatos, que estabelece produção própria ou de terceiros, normalização, escolha do local de acesso, e escolha do nível de proteção</td></tr>
<tr><td>de formato, que determina quais os formatos serão permitidos no repositório</td></tr>
<tr><td>de digitalização de documentos, que determina como deverá ser realizada a digitalização de documentos, caso necessário e possível</td></tr>
<tr><td>de normalização de formatos, que determina como deverá ser feito a normalização de formatos caso necessário</td></tr>
<tr><td>de segurança e preservação da informação, que determina como será feita a segurança e a preservação das informações inseridas no repositório</td></tr>
</table>

Processos específicos para Repositórios Digitais (cont.)	
Padronizar conjuntos de metadados, possibilitando escolher tipos de metadados	
Possibilitar o auto-arquivamento	Oferecer avisos de submissões recentes e isso pode ser feito na página principal do repositório, nas páginas personalizadas dos usuários ou em envio de e-mails
	Contribuir para avaliação dos trabalhos científicos por meio de convites e incentivos aos usuários para participar nas avaliações
Possibilitar heterogeneidade de formatos de arquivos	
Permitir publicação e inclusão de documentos em vários locais, permitindo duplicar os documentos em diversos ambientes de informação e garantindo o direito autoral	
Oferecer gerenciamento de versões do texto, que possibilita as atualizações nos textos depositados	
Oferecer mecanismo de discussão e avaliação entre os pares	
Disponibilizar recurso de classificação de pareceres, que possibilita a recuperação de documentos que foram ou não analisados pelos pares, além dos mesmos poder classificar os documentos analisados	
Disponibilizar relações de documentos com coleções, possibilitando que usuários adicionem documentos em suas coleções pessoais	
Disponibilizar citações relacionadas ao documento, apresentando as citações dos documentos realizadas por outros documentos	
Possibilitar a inserção de comentários e sugestões, oferecendo um espaço de acesso para que o usuário	
Oferecer processo automático de comentários, possibilitando acesso e direcionamento de comentários	
Oferecer serviço ou boletim de alerta por meio de um serviço que gerencia mensagens relativas ao acesso e ao auto-arquivamento	
Oferecer serviço de documentação e catálogo bibliográfico para auxiliar no acesso a documentos	
Oferecer um fórum de acesso livre como um local de debate em que todos os usuários podem acessar	
Possibilitar interoperabilidade, permitindo troca de informações com outros sistemas por meio de protocolos	
Possibilitar preservação da informação	Migração – transferência de conteúdo de um suporte a outro, por meio de gerenciamento de versões
	Emulação – simulação de um *software* compatível com o arquivo
	Preservação da tecnologia/suporte – mantém várias versões de suporte

Processos específicos para Repositórios Digitais (cont.)
Demonstrar a visibilidade da instituição, que pode além de disponibilizar os trabalhos científicos, oferecer indicadores da produção da instituição
Oferecer segurança da informação, utilizando padronização com timbre da instituição nos documentos, além de outros elementos como criptografia, marca d'água, certificação digital e legenda bibliográfica
Oferecer informações sobre: repositório, instituição, mapa do site, divulgação, parcerias, treinamento, instruções, manutenção e atualização, help, experiências de outras instituições, *softwares* e tecnologias etc.

Quadro 1 – Processos específicos para repositórios digitais

PROCESSOS ESPECÍFICOS DE ACESSIBILIDADE

São apresentados no quadro 2 alguns processos específicos de acessibilidade que podem potencializar os repositórios digitais, aumentando as opções de acesso ao mesmo. Esses processos são baseados nos princípios de acessibilidade encontrados no WCAG 2.0[16] (W3C Working Draft 17 May 2007), que é um documento que explica como tornar o conteúdo Web acessível para pessoas com diferentes condições sensoriais, linguísticas e motoras.

Processos específicos de Acessibilidade
Fornecer alternativas de não-texto de modo que possa ser mudado para outro tipo como sons, símbolos ou uma linguagem mais simples, garantindo a transformação harmoniosa da informação.
Fornecer alternativas sincronizadas para multimídia, utilizando recursos que funcionem em ambientes multimídia como texto, som e vídeo
Criar várias maneiras de apresentação do índice como em forma de texto, vídeo e/ou LIBRAS
Oferecer mais de um tipo de mapa do site podendo auxiliar o usuário em relação a navegação do site
Utilizar toda a funcionalidade disponível do teclado como ajuda aos usuários com deficiência motora e visual na utilização de *softwares* de leitura
Incentivar o uso de índice e oferecer bastante tempo para ler para facilitar o processo de busca da informação
Oferecer mapa do site específico para usuários com necessidades

16 Disponível em: <http://www.w3.org/WAI/>

Processos específicos de Acessibilidade (cont.)
Fazer o texto legível e compreensível a todos, utilizando uma linguagem objetiva e clara
Fazer as páginas que possam manter a configuração em diferentes *browsers* e informar o usuário sobre a configuração mínima necessária para uma boa visualização
Oferecer ajuda para usuários evitarem erros
Maximizar a compatibilidade com agentes atuais e futuros, incluindo as tecnologias assistivas como *software* de leitor de tela e de conversão para Linguagens de Sinais (ex. LIBRAS)
Oferecer opção de modificação de tamanho de fonte para auxiliar na leitura de usuários que possuem alguma dificuldade visual, aumentando e diminuindo o tamanho da letra
Oferecer opção de modificação de fundo da página (contraste) para auxiliar na leitura de usuários que possuem sensibilidade ao monitor, podendo modificar a cor do fundo da página para escuro ou claro

Quadro 2 – Processos específicos de acessibilidade

PROCESSOS ESPECÍFICOS DE USABILIDADE

Alguns processos específicos de usabilidade são apresentados no quadro 3 com o intuito de aumentar a utilização dos repositórios digitais pela comunidade de usuários. Esses processos foram baseados nas diretrizes apresentadas por Nielsen e Tahir (2002), as quais foram selecionadas em consequência da relevância de suas utilidades e também para reduzir a redundância de algumas recomendações que são muito semelhantes, ou seja, elas foram selecionadas para representar um conjunto de processos mais reduzido e expressivo e menos subjetivo, minimizando a redundância na avaliação/análise dos ambientes.

Processos específicos de Usabilidade
Exibir o nome da empresa e/ou logotipo ou slogan para garantir a confiabilidade da instituição
Enfatizar as tarefas de mais alta prioridade, disponibilizando na página principal as informações mais relevantes
Agrupar informações da empresa por assunto, disponibilizando de maneira lógica e agrupada as informação do repositório e da instituição
Incluir um link da *homepage* para uma seção "Sobre Nós", apresentando informações sobre a instituição

Processos específicos de Usabilidade (cont.)
Incluir um link "Fale Conosco", apresentando uma forma do usuário entrar em contacto com a instituição ou responsável do ambiente
Usar seções e categorias de rótulo, usando a linguagem do cliente
Evitar conteúdo redundante, utilizando textos que não cansem o usuário e repetindo a mesma informação
Não utilizar frases eruditas nem dialeto, utilizando linguagem clara e simples
Empregar padrões e estilo com consistência como letra maiúscula, cores, fontes etc
Explicar o significado de abreviações, explicando as siglas
Facilitar o acesso aos itens apresentados recentemente na homepage, reservando um espaço para informações recentes
Diferenciar links e torná-los fáceis de serem visualizados, padronizando os links
Permitir links coloridos para indicar os estados visitados e não visitados
Usar nomes de links significativos
Disponibilizar para os usuários uma caixa de entrada na homepage para inserir consultas de pesquisa
Não oferecer um recurso para "pesquisar na web", na função de pesquisa do site
Rotular gráficos e fotos se os respectivos significados não estiverem claros para esclarecer as informações das ilustrações
Permitir que os usuários decidam se desejam ver uma introdução animada de seu site, não deixar a opção de animação predefinida
Usar texto com contraste e cores de plano de fundo, para que os caracteres fiquem o mais legíveis possível
Evitar a rolagem horizontal, pois os usuários não estão acostumados a utilizarem este tipo de barra, além de que ela possibilita frases muito longas na tela
Usar raramente menus suspensos e tomar cuidado para as informações do menu não sobreporem outras
Incluir uma descrição resumida do site no título da janela de forma significativa
Evitar janelas *pop-up*, que são aquelas que abrem sem o consentimento do usuário
Informar se o website ficar paralisado ou partes importantes do website não estiverem funcionando
Atualizar somente o conteúdo realmente modificado, como as atualizações de notícias
Explicar para os usuários os benefícios e a frequência de publicação, antes de solicitar seus endereços de e-mail, explicando ao usuário o porquê da necessidade das informações específicas dele.
Mostrar aos usuários a hora da última atualização de conteúdo, não a hora atual gerada pelo computador

Processos específicos de Usabilidade (cont.)
Disponibilizar mapa do site, situando o usuário em relação a sua localização e navegação
Diminuir textos longos, deixando as informações claras e objetivas
Não utilizar de forma excessiva as ilustrações e animações, disponibilizando apenas itens de interesse e com objetivos bem definidos
Padronizar as páginas do site, indicando ao usuário que ele navega por um mesmo site
Oferecer *feedback*, possibilitando um retorno ao usuário em relação ao seu desempenho
Possuir poucas propagandas
Oferecer cursor com comportamento padronizado
Dar enfoque no conteúdo e não nos adereços
Possibilitar retorno à página principal
Possibilitar acesso às informações por meio de poucos comandos
Utilizar mensagens de erro com vocabulário neutro, não utilizando códigos que os usuários não conseguem entender
Evitar caracteres especiais, não utilizando símbolos ou caracteres que usuários pode não reconhecer

Quadro 3 – Processos específicos de usabilidade

PROCESSOS ESPECÍFICOS DE QUALIDADE

O quadro 4 apresenta processos específicos de qualidade que podem aumentar o desempenho, confiabilidade e outras características que envolvem o comportamento do ambiente. Alguns desses processos foram baseados em normas de qualidade de *software*, principalmente a ISO 9000, que aborda um consenso sobre as características essenciais para garantir a qualidade de sistemas de informação.

Processos específicos de Qualidade
Adequar o uso das funções de acordo com o objetivo do ambiente e público-alvo
Garantir acurácia do sistema, garantindo que os resultados atinjam seus objetivos
Garantir conformidade das funções segundo normas e leis em vigor, verificando normas e decretos em ambiências digitais
Garantir segurança de acesso

Processos específicos de Qualidade (cont.)
Garantir maturidade na confiabilidade, mantendo uma frequência baixa de falhas
Garantir tolerância a falhas mantendo o desempenho, mesmo quando ocorre um erro interno
Garantir recuperabilidade mesmo em caso de falhas
Garantir inteligibilidade, minimizando esforço para o usuário compreender o sistema
Garantir apreensibilidade, minimizando esforço para o usuário aprender a usar o *software*
Garantir operacionalidade, minimizando esforço para o usuário operar o *software*
Oferecer comportamento tempo-resposta, oferecendo uma curta resposta do sistema ao usuário
Oferecer comportamento dos recursos, estabelecendo quantidades de recursos
Garantir manutenibilidade, identificando problemas, removendo-os e adaptando a mudanças, bem como identificando evidências sobre os riscos de efeitos inesperados em caso de mudanças e executando testes em caso de modificações
Oferecer adaptabilidade para verificar a capacidade de adaptar-se a ambientes diferentes
Garantir capacidade para ser instalado
Garantir conformidade quanto à portabilidade, identificando atributos do *software* que identificam o nível de padronização no que se refere à portabilidade

Quadro 4 – Processos específicos de qualidade

PROCESSOS ESPECÍFICOS DE PERSONALIZAÇÃO E CUSTOMIZAÇÃO

Alguns processos relacionados à personalização e customização de interfaces e conteúdos estão apresentados no quadro 5 a fim de aumentar a interatividade dos usuários nos repositórios digitais. Esses processos foram identificados por meio de uma análise realizada em vários tipos de ambientes informacionais digitais, por meio de observação direta não-participante.

Processos específicos de Personalização e Customização	
Possibilitar escolha de cor	do fundo da página
	da letra do título
	da letra do texto
	da letra dos links
	dos botões

Processos específicos de Personalização e Customização (cont.)	
Possibilitar alterar o tamanho	da fonte do texto em geral
	da fonte selecionada
	de imagens
Possibilitar alterar o tipo	da fonte do título
	da fonte do texto
	da fonte do link
Escolher barra de navegação, permitindo ao usuário escolher visualizar a página com ou sem barra de navegação.	
Escolher quantidade de itens por busca, determinando quantos itens devem ser visualizados no resultado da busca	
Escolher barra de rolagem, permitindo navegar verticalmente ou horizontalmente	
Escolher formas de apresentação do resultado da busca, permitindo ao usuário escolher a forma de apresentação das informações recuperadas, como: em tabela, em lista com descrição e em ícones	
Inserir em espaço (mural) livre por meio de um espaço para que o usuário possa inserir alguma informação que deve ficar na página do website	
Oferecer espaço de sugestão por meio de um espaço para que o usuário possa inserir alguma informação que deve ser encaminhada para o responsável do website	
Escolher idioma	das informações das páginas
	das informações dos documentos
Sugerir outros termos de busca. Isso é muito utilizado em bibliotecas digitais ou em sites de buscas quando o usuário digita alguma palavra de seu interesse que pode ter muitos significados ou pode haver outros assuntos similares. O site apresenta várias outras opções de termos, as quais podem lhe auxiliar na busca	
Reservar espaço para discussão e opiniões por meio de um espaço para que o usuário possa inserir informações, contudo é necessário o cadastro de usuários, determinando participação em comunidades, pois essa técnica é destinada para discussão e debate de temas específicos	
Oferecer opção de metadados, que pode tanto oferecer resultados de buscas por meio de algum tipo de metadados (por exemplo, pode-se visualizar o resultado de uma busca tanto em uma linguagem normal em forma de lista de links como também em uma opção de MARC21) quanto permitir ao usuário inserir informações nos campos dos metadados determinados	
Identificar uma informação já acessada, possibilitando ao usuário informar ao website se o mesmo já possui determinada informação/produto ou se ele não tem interesse nela	
Recomendar informações ao usuário	Por meio da filtragem colaborativa automática, que utiliza o conceito de similaridade entre os usuários
	Por meio da filtragem baseada em conteúdo, que utiliza a similaridade entre os produtos
	Por meio de frames de recomendação, que utilizam relações entre produtos, indicando com que probabilidade um produto pode ser consumido caso o usuário esteja comprando outro

Processos específicos de Personalização e Customização (cont.)	
Recomendar informações ao usuário	Por meio de sistemas de recomendação híbridos, que é a utilização das duas primeiras técnicas simultaneamente: a filtragem colaborativa e a filtragem baseada em conteúdo, em que o sistema faz relações entre usuário e entre usuário e produto
Utilizar funções como:	Adicionar, que permite adicionar uma informação/produto em sua página principal
	Salvar, que permite salvar uma informação ou a própria página
	Excluir, que permite excluir uma informação/produto em sua página principal
	Acessar/buscar, que permite ao usuário realizar buscas;
	Enviar por e-mail, que permite ao usuário enviar informação por e-mail
	Imprimir, que permite ao usuário a impressão de informações ou da própria página
	Ordenar, que permite ao usuário ordenar tanto os resultados de busca, quanto as informações armazenadas em sua página pessoal
	Receber de informações por e-mail de interesses pessoais, que possibilita várias opções ao usuário para que ele possa escolher as informações que ele deseja receber por seu e-mail
	Indicar para um amigo, que permite ao usuário indicar informações ou o próprio website para amigos
Permitir a montagem de carteirinha com personagens, geralmente utilizada em sites infantis, em que o usuário pode escolher algum personagem preferido e uma carteirinha é montada para ser utilizada como sua identificação no site	
Visualizar apenas itens selecionados para acessar de forma mais rápida apenas alguns documentos de um determinado assunto por exemplo	
Determinar prioridades, que possibilita ao usuário determinar qual informação/produto é mais relevante para ele	

Quadro 5 – Processos específicos de personalização e customização

Nos quadros 1, 2, 3, 4 e 5 foram apresentados alguns processos que podem auxiliar os arquitetos da informação no desenvolvimento de repositórios digitais com o intuito de aumentar e melhorar a utilização dos mesmos pelos usuários finais. Esses processos podem ser considerados como diretrizes ou recomendações e também podem ser utilizados como check-list para verificação e avaliação.

CONSIDERAÇÕES FINAIS

Apresentou-se aqui uma Arquitetura da Informação que pode ser aplicada para o desenvolvimento e avaliação de Repositórios Digitais a fim de melhorar a qualidade de forma geral, envolvendo vários assuntos como política, interoperabilidade, metadados etc., bem como processos específicos para repositórios digitais, de acessibilidade, usabilidade, qualidade, personalização e customização.

Para a elaboração dessa arquitetura da informação foi realizada uma análise em vários tipos de ambientes científicos digitais, principalmente nos repositórios digitais, que são tema central deste texto, com o propósito de identificar características e recursos específicos. Esse tipo de ambiente é relevante para as comunidades de usuários em consequência da sua crescente utilização para gerenciamento da produção institucional e por apresentarem benefícios efetivos como auxílio na aquisição e construção de trabalhos científicos, no acesso irrestrito, nos intercâmbios e na troca de informações.

Concluindo, espera-se que a Arquitetura da Informação proposta possa auxiliar no desenvolvimento dos repositórios digitais oferecendo recursos interativos e otimizando assim a interação do usuário com o sistema, além de oferecer serviços de coleta de dados, de personalização das informações do usuário, de divulgação do ambiente e de documentação em geral. A Arquitetura da Informação pode ainda permitir um planejamento prévio dos serviços e do funcionamento do ambiente informacional, conduzindo o desenvolvedor por meio de diretrizes e processos e auxiliando na construção de um ambiente com características e políticas próprias e específicas.

REFERÊNCIAS

ABNT- ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **NBR ISO 9000/2000: sistema de gestão da qualidade: fundamentos e vocabulário**. Rio de Janeiro, 2004. Disponível em: < http://www.scribd.com/doc/27778803/NBR-ISO-9000-Sistemas-de-Gestao-da-Qualidade-Fundamentos-e-Vocabulario>. Acesso em: 12 dez. 2009.

ALVES, R. C. V. **Metadados como elementos do processo de catalogação.** 2010. 132 f. Tese (Doutorado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.

BARANAUSKAS, M. C. C; MANTOAN, M. T. E. Acessibilidade em ambientes educacionais: para além das guidelines. **Revista *On-line* da Bibl. Prof. Joel Martins**, São Paulo,v. 2, n. 2, p. 13-22, 2001.

BOERES, S. A; MÁRDERO ARELLANO, M. A. Políticas e estratégias de preservação de documentos digitais. In: CINFORM – ENCONTRO DE CIÊNCIA DA INFORMAÇÃO, VI , 2005, Salvador. **Anais...** Salvador: UFBA, 2005. Disponível em: <http://dici.ibict.br/archive/00000263/01/Preserva%C3%A7%C3%A3o_VI_CINFORM.pdf>. Acesso em: 02 fev. 2010.

CAFÉ, L. et al. Repositórios institucionais: nova estratégia para publicação científica na Rede. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 26., 2003, Belo Horizonte. **Anais eletrônicos...** Belo Horizonte: INTERCOM – Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, 2003.

CAMARGO, L. S. A. **Metodologia de desenvolvimento de ambientes informacionais digitais a partir dos princípios da arquitetura da informação**. 2010. 287 f. Tese (Doutorado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.

COELHO, C. **Um repositório digital para a U.Porto**: Relatório preliminar. Universidade do Porto, Biblioteca Virtual. 2005. Disponível em: < http://repositorio-aberto.up.pt/handle/10216/6701 >. Acesso em: 02 fev. 2010.

CORRADI, J. A. M; VIDOTTI, S. A. B. G. Elementos de acessibilidade em ambientes informacionais digitais: bibliotecas digitais e inclusão social. IN: SEMINÁRIO INTERNACIONAL DE BIBLIOTECAS DIGITAIS BRASIL, 2007, São Paulo. **Anais...** São Paulo, 2007. Disponível em: < http://www.sibi.usp.br/sibi/sibdb/trab_aprov.htm>. Acesso em: 20 out. 2009.

CROW, R. **The case for institutional repositories**: a SPARC position paper. The Scholarly Publishing & Academic Resources Coalition 21 Dupont Circle. Washington, DC 2002. Disponível em: <http://www.arl.org/sparc/bm~doc/ir_final_release_102.pdf>. Acesso em: 15 dez. 2008.

FREIRE, A. P.; FORTES, R. P. M. Avaliação e re-engenharia da interface de uma aplicação Web de acordo com normas de acessibilidade. In: SIMPÓSIO SOBRE FATORES HUMANOS EM SISTEMA COMPUTACIONAIS, VI , 2004, Ribeirão Preto, **Anais...** Ribeirão Preto, 2004, p. 181-184. Disponível:<http://safe.icmc.usp.br:9673/safe/scientific-production/FreireFortesWIC.pdf>. Acesso em: 25 out. 2008.

GILLILAND-SWETLAND, A. J. La definición de los metadatos. In: BACA, M.; GILL, T. Introducción a los metadatos: vías a la información digital. [S. l.]: GETTY, 1999. p. 1-9.

LATHAM, D. Information architecture: notes toward a new curriculum.

Journal of the American Society for Information Science and Technology, v. 53, n. 10, p. 824-830, 2002. Disponível em: <http://www.eric.ed.gov/ERICWebPortal/custom/portlets/recordDetails/detailmini.jsp?_nfpb=true&_&ERICExtSearch_SearchValue_0=EJ654158&ERICExtSearch_SearchType_0=eric_accno&accno=EJ654158>. Acesso em: 02 fev. 2010.

McGEE, J.; PRUSAK, L. **Gerenciamento estratégico da informação**. Tradução de Astrid Beatriz de Figueiredo. Rio de Janeiro: Campus, 1994. p. 129-149.

MORVILLE, P.; ROSENFELD, L. **Information architecture for the World Wide Web.** Sebastopol, CA: O'Reilly, 2006. 504p.

NIELSEN, J.; TAHIR, M. ***Homepage*: usabilidade - 50 *Websites* desconstruídos**. Tradução de Tereza Cristina Félix de Souza. Rio de Janeiro: Campus, 2002. 315p. Disponível em: < http://www.ime.usp.br/~anderson/novatec.html>. Acesso em: 02 fev. 2010.

ORTIZ-REPISO JIMÉNEZ, V. Qué enseñamos después del MARC?. **Organizacion Del Conocimiento en Sistemas de Información y Documentación**, Zaragoza, v. 03, p. 217-225, 1999.

RODRIGUES, E. **Acesso livre à literatura científica**: o repositório institucional e a política de auto-arquivo da Universidade do Minho. 2005. Disponível em: < https://repositorium.sdum.uminho.pt/handle/1822/3478>. Acesso em: 25 jan.2009.

ROSENFELD, L; MORVILLE, P. **Information architecture for the World Wide Web.** Sebastopol, CA: O'Reilly, 1998. 202p.

SANTOS, R . **Usabilidade de interfaces e arquitetura de informação**: alguns aspectos da organização de conteúdo para o meio digital. In: CONGRESSO DA BRASILEIRO DE ERGONOMIA, XI . CONGRESSO LATINO AMERICANO DE ERGONOMIA, VI. ENCONTRO ÁFRICA-BRASIL DE ERGONOMIA, III. FÓRUM SUL BRASILEIRO DE ERGONOMIA, III, 2001, Gramado. **Anais eletrônicos...** Gramado, 2001. 6p. Disponível em: http://www.robsonsantos.com/trabalhos/usab_ai.PDF. Acesso em: 26 mar. 2009.

SARMENTO, F. et al. **Algumas considerações sobre as principais declarações que suportam o movimento acesso livre**. 2005. Disponível em: <http://repositorium.sdum.uminho.pt/bitstream/1822/4282/1/Sarmento+Miranda+ Baptista+Ramos+-+Vers%C3%A30+Final.pdf>. Acesso em: 30 mar. 2009.

SAYÃO, L. F. Preservação de revistas eletrônicas. In: FERREIRA, S. M. S. P; TARGINO, M.G (Org.). **Mais sobre revistas científicas**: em foco a gestão. São Paulo: Editora Senac; Cengage Learning, 2008.

SILVINO, A. M. D; ABRAHÃO, J. I. Navegabilidade e inclusão digital: usabilidade e competência. **RAE-eletrônica**, v. 2, n. 2, jul./dez. 2003. Disponível em: <http://www.rae.com.br/artigos/1808.pdf>. Acesso em: 02 fev. 2010.

SIQUEIRA, A. H. de. **A lógica e a linguagem como fundamentos da arquitetura da informação**. 2008. 143 f. Dissertação (Mestrado em Ciência da Informação e Documentação) – Universidade de Brasília. Disponível em: <http://bdtd.bce.unb.br/tedesimplificado/tde_busca/arquivo.php?codArquivo=3180> Acesso em: 02 fev. 2010.

SOTILLOS SANZ, L. La arquitectura de la información en los medios digitales. In: KNAPP BJERÉN, A. (Coord.). **La experiencia del usuario**. Madrid: Anaya Multimedia, 2002.

TORRES, E. F.; MAZZONI, A. M.; ALVES, J. B. M. A acessibilidade à informação no espaço digital. **Ciência da Informação**, Brasília, v. 31, n. 3, p. 83-91, set./dez. 2002.

VECHIATO, F. L. **Repositório digital como ambiente de inclusão digital e social para usuários idosos**. 2010. 183 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília.

VIANA, C. L. M; MÁRDERO ARELLANO, M. A; SHINTAKU, M. **Repositórios institucionais em ciência e tecnologia**: uma experiência de customização do DSpace. In: SIMPÓSIO DE BIBLIOTECAS DIGITAIS, 3., 2006, São Paulo. **Anais eletrônicos...** São Paulo: CRUESP, 2005. Disponível em: http://bibliotecas-cruesp.usp.br/3sibd/docs/viana358.pdf. Acesso em: 28 mar. 2009.

VIDOTTI, S. A. B. G; CUSIN, C. A.; CORRADI, J. A. M. **Acessibilidade digital sob o prisma da Arquitetura da Informação**. In: GUIMARÃES, J. A. C.; FUJITA, M. S. L. Ensino e pesquisa em Biblioteconomia no Brasil: a emergência de um novo olhar. São Paulo: Cultura Acadêmica, 2008.

WATERS, D; GARRETT, J. **Preserving digital information: report of the task force on archiving of digital information**. Washington, DC: The Commission on Preservation and Access; The Research Library Group, 1996. p. 1-71. Disponível em: < http://www.clir.org/pubs/reports/pub63watersgarrett.pdf>. Acesso em: 03 mai. 2010.

WINCKLER, M; PIMENTA, M. S. Avaliação de usabilidade de Sites Web. In: ESCOLA REGIONAL DE INFORMÁTICA, 2002, Porto Alegre. **Anais...** Porto Alegre: SBC, 2002. Disponível em: <http://www.funtec.org.ar/usabilidadsitiosWeb.pdf>. Acesso em: 02 fev. 2010.

WURMAN, R. S. **Information architects**. Zurich: Graphis, 1996.

WURMAN, R. S. **Information anxiety 2.** Indianapolis: QUE, 2001.

WORLD WIDE WEB CONSORTIUM. *Essencial components of Web accessibility,* Madison, 2007. Disponível em: <http://www.w3.org/WAI/intro/components.php>. Acesso em: 02 fev. 2010.

Repositórios institucionais:
por uma memória criativa no ciberespaço

Vera Dodebei

MEMÓRIA E CONHECIMENTO

Memória e conhecimento foram, por muitas vezes, considerados conceitos idênticos, tal era a disputa pelo domínio e uso da informação. Com a mudança do meio de transmissão do conhecimento de oral para registros visuais (aí inserida a escrita), as sociedades passam a enfrentar um novo desafio em decorrência da tomada de consciência das dimensões do tempo e do espaço envolvidas nas operações de transmissão da memória ou do conhecimento.

A discussão que se segue sobre os meios de acumulação de conhecimentos e memórias tem como objetivo criar uma linha de fuga, em direção ao futuro, que nos permita avançar no processo comunicacional sem ficarmos atrelados a uma determinada era tecnológica. Já sabemos que não há como guardar toda a produção do conhecimento e sabemos também que o custo da preservação de memórias é alto, o que nos obriga a fazer escolhas. Com o advento da criação do ciberespaço e a possibilidade tecnológica de criação de memórias eletrônicas, parece-nos que não há mais como adiar o exame dessas questões, sob pena de tombarmos ou para o lado memorialista da proteção absoluta de todas as informações e suas variantes, ou para o lado da transformação constante e sem rastros, restos ou ruínas de todo o conhecimento produzido na sociedade, a exemplo do que ocorria na era da oralidade mítica[1].

Dentre as mais diversas formas e meios de memória, este texto visa discutir uma pequena parte desse universo de informações que circu-

1 Sobre a discussão das eras tecnológicas e o estatuto da Memória Social ver (DODEBEI, 2000)

lam nas sociedades. Nossa atenção é para com a memória científica de instituições que visam garantir a produção, circulação e guarda de informações em meio digital. Em síntese, pretendemos fazer uma análise da evolução dos meios de comunicação científica, com foco no papel que os Repositórios Institucionais representam à questão, sob o viés das teorias e dos estudos contemporâneos no campo da Memória Social.

Organizamos nossa exposição em quatro momentos. Inicialmente, discutiremos os meios de memória utilizados pelas sociedades com o objetivo de transmitir a experiência vivida; o segundo momento é dedicado à análise das memórias eletrônicas com destaque para os repositórios institucionais; no terceiro, apresentamos uma reflexão sobre a compreensão dos modelos criados para explicar o processo de transferência de informações na sociedade; e o quarto momento é dedicado a projetar uma linha de fuga entre as operações de proteção e de disseminação da informação, o que denominamos de memória criativa no ciberespaço.

É importante ressaltar que a questão que nos instiga não está relacionada ao modo de constituição dos repositórios institucionais, nem tampouco às tecnologias postas em prática para sua operacionalização. Nossa discussão visa subsidiar o entendimento deste recurso como um meio de memória no mundo contemporâneo. Compreender os mecanismos da memória pode nos ajudar a escolher entre procedimentos de preservação que visem ou a manutenção da totalidade da oferta de informações (acúmulo de informações), ou a seleção destas (disputa entre o que lembrar e o que esquecer).

MEIOS DE MEMÓRIA

A arte da memória, criada pelos gregos, foi tema incandescente no passado e teve como defensor da técnica de memorizar ou imprimir lugares e imagens na memória o renascentista Giordano Bruno. Frances Yates (2007) nos diz que a arte da memória é como uma escrita interior. Se a partir do momento em que conhecemos as letras do

alfabeto podemos escrever o que é ditado e ler o que escrevemos, do mesmo modo, aqueles que aprenderam a mnemônica podem colocar em lugares específicos aquilo que ouviram falar de memória. No entanto, mesmo que sejamos muito treinados a criar esses "lugares" na memória, o fato é que estamos permanentemente realocando esses espaços, quer dizer, estamos processando continuamente as séries de lugares em nossa memória individual.

A importância do exercício de memória individual na antiguidade levou filósofos, clérigos e cientistas a disputas de poder que, não raro, terminavam em morte. Muitos desses estudiosos da memória foram acusados de magia e incluídos na lista da Santa Inquisição, como ocorreu a Giordano Bruno que morreu queimado no ano de 1600. A arte da memória ou mnemotécnica representava um sistema de imagens que poderiam levar a fixar na memória coisas/conceitos (*res*) e palavras (*verba*). Segundo Yates (2007, p. 25), "isto quer dizer que a 'memória para coisas' cria imagens para nos lembrarmos de um argumento, de uma noção, ou de uma coisa, e a "memória para palavras" busca imagens para que nos recordemos de cada palavra." Parece que o objetivo final era o de memorizar o mundo e organizar o discurso, como explicitado nas cinco partes do exercício da retórica: coisas verdadeiras; organização dessas coisas; atribuição de palavras às coisas; memorização das coisas e das palavras; comunicação das coisas por palavras.

Do ponto de vista da organização dessas imagens no cérebro, ensinava-se a criar lugares como um edifício dividido em cômodos, onde se alojaria em cada um dos cômodos a imagem de coisas e de palavras. Nesse sentido, o poder da memória residia na faculdade de selecionar o maior número de informações arquitetadas, de forma organizada, para que cada uma delas pudesse ser recuperada pela memória no momento necessário. Os mapas visuais criados por Giordano Bruno e outros como os de Raimundo Llulos certamente inspiraram as classificações filosóficas e posteriormente as classificações bibliográficas que não tinham outro sentido que o de organizar a memória do conhecimento humano.

O exercício de memória individual, na Idade Média, dá lugar à criação de sistemas arquivísticos e bibliográficos de organização dos

meios escritos de memória coletiva. A biblioteca de Warburg descrita por Settis (2000), o sistema de organização conceitual para ajudar a memória de escritores, criado por Peter Mark Roget (1998) no início do século XIX na Inglaterra, as classificações bibliográficas de Melvil Dewey, Henry Evelyn Bliss e Shiyali Ramamrita Ranganathan na passagem do século XIX para o século XX (PIEDADE, 1983), sem mencionar o espírito enciclopedista[2] são exemplos de criação de memórias auxiliares ao conhecimento registrado e acumulado em coleções de livros, papéis e objetos. Enfim, os novos meios de memória eletrônica que aparecem nas últimas décadas do século XX nos convidam a pensar a memória no século XXI, suas configurações e desafios.

Esse olhar para o passado nos conduziu à época anterior ao surgimento da era da tecnologia da escrita, quando a transmissão dos saberes era feita de forma oral, sem a preocupação com registros em suportes materiais, daí a importância do exercício de nossa memória individual. Esses dois meios de memória (a oralidade e a escrita) coexistem nos dias atuais e entram em disputa com novas mídias digitais no espaço virtual. À memória individual, que ao invés de acumular processa para produzir novas informações, e à memória auxiliar, criada pela sociedade da escrita representada principalmente por arquivos, bibliotecas e museus, se junta a memória eletrônica que transforma o visual e o oral em números. Em que medida essas três formas de memória interagem e podem nos ajudar a compreender os mecanismos de proteção e de disseminação dos saberes na sociedade contemporânea?

Le Goff (2003, p. 422) nos diz que o estudo da memória social é um dos meios fundamentais de abordar os problemas do tempo e da história, relativamente aos quais a memória está ora em retraimento, ora em transbordamento. Os tempos da memória, para Le Goff, são marcados também pela transição entre a oralidade e a escrita e ele a periodiza em cinco segmentos: a memória étnica nas sociedades sem escrita, ditas selvagens[3]; o desenvolvimento da memória, da oralidade

2 Cf. Mapa conceitual para a estrutura dos volumes da Enciclopédia Einaudi, por exemplo, o volume sobre Cultura Material (BUCAILLE; PESEZ, 1989).

3 Menção ao que Jack Goody chama de "a domesticação do pensamento selvagem" (LE GOFF, 2003).

à escrita, da Pré-história à Antiguidade; a memória medieval, em equilíbrio entre o oral e o escrito; os progressos da memória escrita, do século XVI aos nossos dias; e, os desenvolvimentos atuais da memória.

Para esse último período, o da memória no último século, Le Goff (2003, p. 463) se inspira no que Leroi-Gourhan denominou de "memória em expansão". No entanto, ele não discute essa memória em expansão e apenas aponta que: "os desenvolvimentos da memória no século XX, sobretudo depois de 1950, constituem uma verdadeira revolução da memória, e a memória eletrônica não é senão um elemento, sem dúvida o mais espetacular." Segundo o autor, a tensão entre oralidade e escrita conduz os estudos históricos sobre a sociedade, e o aparecimento da escrita na sociedade é o que vai transformar profundamente a memória coletiva do mundo oral.

A nova era da produção de registros do conhecimento que tem como marco, no Ocidente, a invenção da imprensa por Gutenberg trouxe um enorme avanço aos meios de comunicação, pois a mobilidade dos impressos no tempo e no espaço permite que a informação seja disseminada a um número maior de pessoas distantes geográfica e temporalmente do emissor da mensagem registrada. Com essa situação, nos habituamos a adquirir registros e a acumulá-los individual e coletivamente.

Em contrapartida ao desejo de acumular ou reter, as ações de mudar, renovar e inovar são também formas de esquecimento igualmente próprios das pessoas e, consequentemente das instituições. Portanto, escolher o que guardar e o que descartar sempre foi parte das atividades de instituições de memória ou lugares de memória, como discute Pierre Nora (1993). Se não há mais meios de memória espontânea, o que era próprio da oralidade, há que se criar lugares para conservar lembranças do passado. Os lugares de memória para Nora não se restringem a espaços físicos apenas, mas se constituem em espaços de ação, de trocas, de comemoração, de manutenção da tradição, onde as escolhas marcam sua presença. Nesse sentido, podemos considerar o ciberespaço como um lugar de memória, sujeito aos embates entre lembrar e esquecer.

ORGANIZAÇÃO DA MEMÓRIA INSTITUCIONAL: BIBLIOTECAS E ARQUIVOS DIGITAIS

As instituições de memória sempre tiveram a preocupação com a seleção dos objetos de suas coleções, ainda que essa seleção fosse de natureza política, como as escolhas o são. O interesse das bibliotecas, dos arquivos, dos museus foi e é o de criar coleções que possam simbolizar o conhecimento acumulado, talvez visando a um coletivo que transcenda à singularidade da produção intelectual. Assim como o mapa mental do conhecimento utilizado por Giordano Bruno, o sonho ou a utopia de reunir em um espaço controlável toda a memória do mundo não é privilégio das sociedades contemporâneas (LEVIE, 2006).

Nesse sentido, no momento em que se discute a constituição de uma memória coletiva (HALBWACHS, 2004), de uma inteligência coletiva (LEVY, 2003), de um arquivo eletrônico coletivo (INTERNET ARCHIVES, 2010) vemos também iniciativas de instituições que buscam, pela seletividade, projetos de valorização dessa produção como patrimônio da humanidade, como é o caso singular da Biblioteca Digital Mundial (2010) e dos Repositórios Institucionais.

BIBLIOTECA DIGITAL MUNDIAL

A Biblioteca Digital Mundial (BDM) é uma coleção de bens culturais mundiais que poderia contar histórias e iluminar os acontecimentos de países e culturas, promovendo a troca e o entendimento multicultural. Ela representa uma mudança em projetos de bibliotecas digitais, do enfoque em quantidade por si só, para qualidade; a quantidade continua a ser uma prioridade, mas não à custa das normas de qualidade estabelecidas durante a fase inicial do projeto de criação do website (BIBLIOTECA DIGITAL MUNDIAL, 2010).

Segundo depoimento de James H. Billington (2009)[4], a BDM não se constitui em uma biblioteca a mais. Ela representa um sítio gratuito na Internet que propõe uma seleção de documentos provenientes de

4 Diretor da Biblioteca do Congresso Americano e principal promotor da BDM.

fundos documentais digitalizados das grandes bibliotecas do mundo. O projeto, de natureza educativa, patrimonial e filantrópica, encampado pela UNESCO em 2005, tem por objetivo reduzir a lacuna digital entre os hemisférios norte e sul, ao tornar acessível um maior número de "documentos fundamentais" que pertencem à cultura planetária, como manuscritos, mapas, filmes.

É Billington (2009) ainda quem afirma que a BDM "absolutamente não compete" com as demais bibliotecas digitais existentes, a exemplo da Europeana (2010)[5] ou Google (2010)[6]. A BDM se difere daquelas, principalmente por sua dimensão multilingue e por sua hiper-seletividade na escolha dos documentos disponibilizados em tempo real. Do acervo da Biblioteca Nacional (do Brasil), por exemplo, foram incorporados à BDM 142 itens relativos à Coleção Thereza Christina[7], registrada na lista do Patrimônio da Humanidade, pela UNESCO e que ilustra a página de abertura da América Latina e Caribe.

Fonte: Biblioteca Digital Mundial (2010)

5 Criada em 2008, a biblioteca digital Europeana dá acesso ao patrimônio cultural e científico. Disponibiliza 6 milhões de itens (EUROPEANA, 2010).

6 Lançada em 2005, a biblioteca digital da Google Inc. disponibiliza 10 milhões de livros (GOOGLE BOOKS, 2010).

7 A Coleção Thereza Christina Maria é composta por 21.742 fotografias organizadas pelo Imperador Pedro II (1825-91) ao longo de sua vida e doada por ele à Biblioteca Nacional do Brasil.

A principal característica da BDM é a seletividade de seu acervo, ou como afirma Billington, a hiper-seletividade. Há muito não encontrávamos na internet projetos que assumissem e defendessem a seleção no momento da oferta de objetos. O modelo de preservação patrimonial que vem sendo mais utilizado até o momento pode ser descrito como o do "excesso"; quer dizer, a prática da multiplicação do objeto lançado na rede para favorecer a sua permanência por um período de tempo maior.

Evidentemente, a seleção, esta palavra-chave já tão conhecida dos processos de organização do conhecimento, implica custos muito elevados. Não foi diferente o que ocorreu com a criação da BDM. O projeto obteve apoio de várias instituições com um aporte de recursos de, aproximadamente, 10 milhões de dólares americanos.

Do ponto de vista da passagem das bibliotecas em suporte papel para o mundo virtual, o esforço de digitalizar o passivo da memória do mundo vem sendo questionado principalmente pelo fator econômico dos direitos autorais. A seletividade com a qual a BDM se moldou, minimiza essa discussão sobre a propriedade intelectual singular, pois transforma os objetos, em sua maioria já de domínio público, em coletivos da humanidade. Diferentemente dessa prática, outras bibliotecas digitais como a Europena e a Google que disponibilizam acervos protegidos pela legislação autoral, enfrentam ainda disputas judiciais para manter seus projetos em atividade na rede.

Podemos supor então que a BDM se classifica no modelo teórico-metodológico da História, no qual o historiador faz a escolha dos documentos que servirão para construir uma versão possível dos acontecimentos do passado, conforme discute Le Goff, em seu texto *História e memória* publicado em 2003. Por oposição, o conjunto de *websites, blogs* e demais redes sociais de iniciativa dos usuários, segue o modelo da memória; imitação, multiplicação, reformatação da informação, são processos caracterizados pelo excesso de informação, têm custos ínfimos e não consideram o perigo da perda dos objetos digitalizados ou criados digitais por valorizarem mais a criação do que a reprodução.

REPOSITÓRIOS INSTITUCIONAIS

Na era digital, os Repositórios Institucionais (RI) representam a memória eletrônica de um grupo de pessoas e, segundo Mark Ware (2004), surgem durante a segunda metade do ano de 2002, a partir da concepção do software DSpace[8] desenvolvido pelo Massachussetts Institute of Technology (MIT), como uma estratégia das universidades para implantar novas rotinas operacionais no processo de comunicação técnico-científica.

A revolução digital afetou o modo como os acadêmicos criam, comunicam e preservam o conhecimento produzido. Com a facilidade de publicação em websites, blogs e outros lugares virtuais os pesquisadores disseminam sua produção sem a preocupação com a proteção dos conteúdos, no longo prazo. Davis e Connolly (2007) indicam que é preciso que as instituições garantam essa preservação. Nesse sentido, os RI assumem essa responsabilidade ao extinguir os direitos autorais das editoras tradicionais e assumir para si a tarefa de garantir o conhecimento produzido por seus pesquisadores. Do ponto de vista das políticas de publicação, eles podem ser vistos como competidores aos meios tradicionais, ou podem ser considerados complementares a eles. O software DSpace utilizado para o desenvolvimento de muitos repositórios, por exemplo, não foi concebido para competir com as editoras tradicionais, mas foi criado como uma tecnologia que permite a reunião de diversos tipos de conteúdos: artigos, teses, materiais culturais, literatura cinzenta, imagens paradas e em movimento, artefatos, entre outros.

Um repositório institucional é uma base de dados digital e virtual (*web-based database*)[9], de caráter coletivo e cumulativo (memória da instituição), de acesso aberto e interoperável que coleta, armazena, dissemina e preserva digitalmente a produção intelectual da instituição.

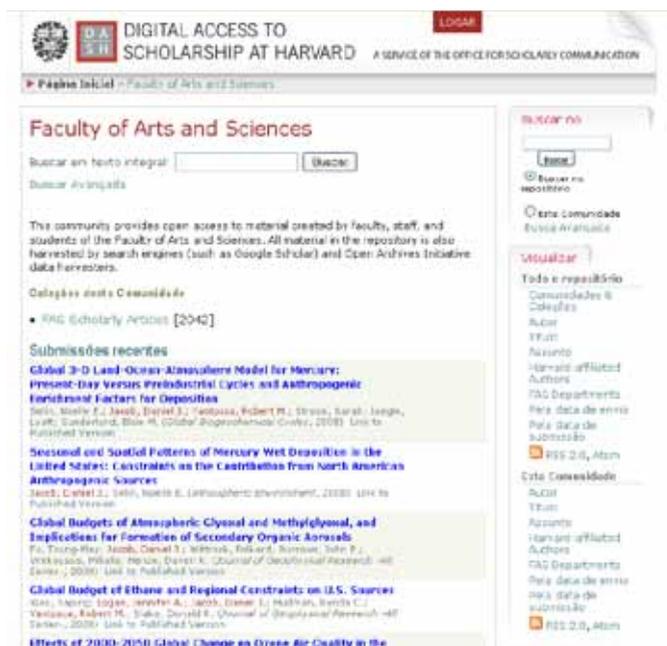
8 DSpace é um software de acesso livre que provê ferramentas para a gestão de recursos eletrônicos e é comumente utilizado com plataforma para os Repositórios Institucionais e sua preservação digital. Desde a sua disponibilização em 2002, ele tem sido instalado em mais de 240 instituições, desde grande universidades até instituições isoladas de ensino e pesquisa. O software é compartilhado sob a licença BSD, o que permite ao seu usuário customizá-lo. (WIKIPEDIA, 2010)

9 Entende-se por *virtual*, o espaço ocupado pelo recurso e *digital*, a sua forma de comunicação.

Do ponto de vista de sua categoria conceitual pode-se dizer que os repositórios são espécies combinadas de arquivo e de biblioteca digitais. A separação feita no passado dessas duas instituições era marcada pela natureza de seu acervo. De um lado, os arquivos representavam a memória das ações administrativas da instituição, incluídos aí os documentos representativos dos credos, políticas, ações, projetos e execução das atividades decorrentes. O recolhimento da documentação era compulsório, o que gerava, no universo analógico, a grandeza de metros e quilômetros de papéis. Já as bibliotecas tinham como objetivos guardar a memória da produção técnico-científica-cultural da sociedade, em equilíbrio entre a oferta e a demanda de informações do grupo para o qual prestava serviços, independentemente do local, dos assuntos tratados e dos meios em que foram produzidos os documentos.

Indiscutivelmente, as ações de seleção se davam tanto nas bibliotecas como nos arquivos e, portanto, ambas as instituições sempre selecionaram suas memórias, fosse ao momento da produção (gestão de documentos para o caso dos arquivos) ou na escolha dos itens que formariam as coleções em uma biblioteca. As justificativas para a seleção eram, de modo geral, voltadas à capacidade de gestão desses documentos, mas mais importante, às políticas de acesso à informação. Uma espécie de filtro pessoal e institucional operava o fluxo de informações, como os embates que ocorrem no interior da consciência em que a luta entre lembrar e esquecer conforma a memória (GONDAR, 2000).

Parece ser de consenso entre os pesquisadores que o principal fator para o desenvolvimento e aceitação dos RI foi a chamada crise da comunicação técnico-científica. O seu cenário pode ser delineado com base em alguns indicadores, entre eles: o alto custo da veiculação de produtos pelas mídias tradicionais; a necessidade de transparência das ações institucionais para subsidiar a avaliação das instituições na luta para a obtenção de recursos; a pressão gerada pela comunidade científica em prol da diminuição da censura, ou proteção excessiva ao acesso às informações; o desenvolvimento acelerado da tecnologia de armazenamento de recursos digitais.

Na universidade de Harvard, em 2008, a Faculdade de Artes e Ciências adotou uma política de acesso aberto à informação que requer que todos os seus autores acadêmicos enviem, automaticamente, uma cópia eletrônica de seus artigos ao repositório digital da universidade. Foi acordado, ainda, que os autores devem conceder automaticamente a permissão dos direitos autorais à universidade para distribuir estes artigos em um sistema de acesso livre.

Fonte: DASH (2010).

No relatório *Open doors and open minds* (2008), os autores comentam que a internet vem proporcionando oportunidades sem paralelo para a expansão do acesso à pesquisa ao derrubar as barreiras econômicas e físicas que impedem o compartilhamento da informação. O ambiente em redes digitais promete democratizar o acesso, reconduzir o conhecimento aos seus nichos originários (colégios invisíveis), acelerar as descobertas, encorajar novos e interdisciplinares enfoques para os mais complexos desafios científicos, e permitir novas estratégias de pesquisa computacional. No entanto, a despeito destas oportunidades

para aumentar o acesso ao conhecimento, os preços dos periódicos continuaram aumentando ao longo das duas últimas décadas, forçando as bibliotecas a cancelarem assinaturas.

Os repositórios institucionais deslocam das editoras de periódicos para a instituição o poder da guarda e do acesso à informação. Além disso, eles quebram o monopólio que as editoras acumularam durante anos com a propriedade dos direitos autorais. Segundo Mark Ware (2004), os repositórios institucionais tiveram importante papel na já mencionada crise da comunicação científica quando se apresentam como indicador tangível da qualidade da universidade, demonstrando a relevância científica, social e econômica de suas pesquisas e aumentando a sua visibilidade e seu valor público.

Mas, retomando a questão da seleção de nossas memórias, resta discutir como esse conhecimento produzido e armazenado nos repositórios de livre acesso vai ser selecionado para compor, de fato, a memória institucional. Examinaremos no próximo item a evolução de alguns modelos de representação do processo de comunicação científica e tecnológica que possam nos indicar uma linha de fuga entre o acúmulo indiscriminado e repetitivo de recursos digitais e a transformação dos recursos em narrativas re-apropriadas e sem autoria.

COMUNICAÇÃO CIENTÍFICA E MEMÓRIA SOCIAL

Selecionamos para o discurso neste item dois modelos de transferência da informação que foram discutidos em trabalho apresentado no 10º Encontro Nacional de Pesquisa em Ciência da Informação em 2009. O segundo modelo, sobre o ciclo da informação em discussão na internet vem sendo objeto de pesquisa no âmbito do projeto *Informação, Memória e Patrimônio: o conceito de acumulação e o ambiente virtual.* A escolha contempla o início do desenvolvimento do tema e considera Frederic Wilfrid Lancaster o marco zero da discussão.

Em seu artigo *The dissemination of scientific and technical information: toward a paperless system,* escrito em 1977 e digitalizado pela University of Illinois Graduate School of Library Science em 2007, Lancaster (2007)

dizia que a informação técnica e científica é transmitida tanto por canais formais como por canais informais. Os canais formais envolveriam o uso de documentos impressos, de vários tipos, enquanto os canais informais diriam respeito à transferência da informação oral. Na representação que constrói indicando a ação dos modelos da oralidade e visualidade (textos e imagens), Lancaster (2007, p. 2, tradução nossa) enfatizava, já naquela época, que "o fluxo de informação, como pode ser visto abaixo, era deliberadamente visto como um ciclo porque o processo é continuo e regenerativo"[10], quer dizer que o conhecimento é continuamente transformado, renovado.

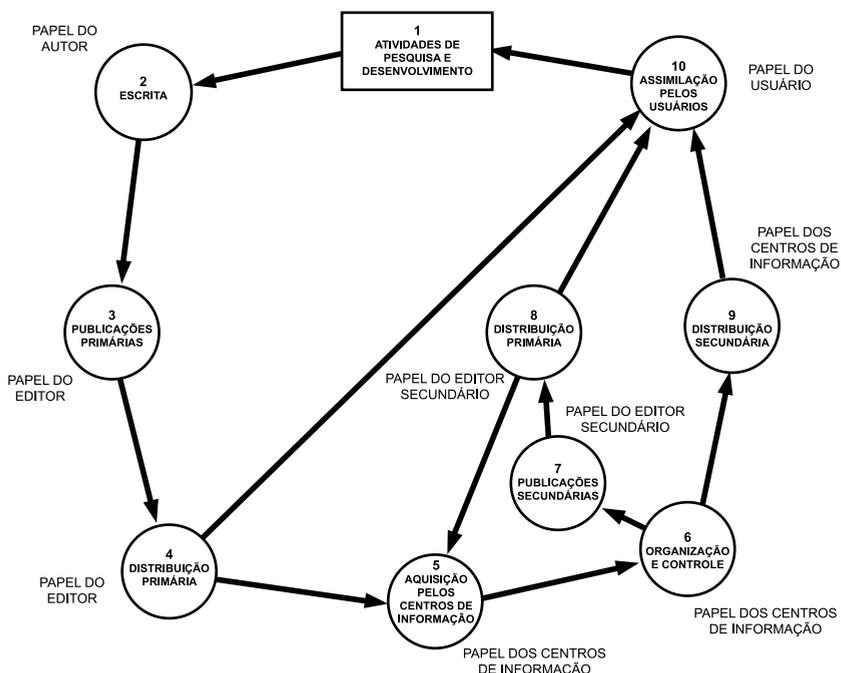



Figura 1 – Disseminação da informação técno-científica por meio de publicações primárias e secundárias
Fonte: Adaptada de Lancaster (1977, p. 3).

Os problemas apontados por Lancaster no ano de 1977 em relação à eficiência da transferência da informação diziam respeito a quatro

10 "The flow of information as depicted in figure 1 is deliberately shown as a cycle because the process is continuous and regenerative."

aspectos principais: 1. a literatura que cresce a taxas exponenciais; 2. a dispersão dessa literatura em mídias diversas; 3. o espaço de tempo cada vez maior entre o resultado da pesquisa e sua divulgação; e, 4. os custos de publicação que ultrapassam os orçamentos do pesquisador individual e das instituições de guarda e disseminação da informação.

Uma das razões apontadas por Lancaster para o aumento dos custos do processo de transferência da informação científica e tecnológica era o fato de que bibliotecas e editoras se encontravam muito distanciadas de outras "indústrias" no que se refere à modernização de seus processos de produção, principalmente no que dizia respeito ao uso da informática. A utilização de computadores no processo de recuperação da informação data de 1950 e, segundo o autor, somente na década seguinte, nos anos 60, as instituições envolvidas no ciclo de transferência da informação começam, de fato, a utilizar os recursos de processamento de dados, ainda no modo remoto (*offline*).

Os efeitos da introdução do processamento de dados nas instituições de memória científica e tecnológica começam a ter visibilidade, nos Estados Unidos, por exemplo, em grandes sistemas de informação como o Defense Documentation Center, a National Aeronautics and Space Administration e a National Library of Medicine. O sistema desenvolvido pela Biblioteca Nacional de Medicina, MEDLARS, que serviu de modelo para muitos países, data de 1964, sendo convertido para acesso em linha (*on-line*) com a sigla MEDLINE, nos últimos anos da década de 1970.

Lancaster considerava que o maior desenvolvimento ocorrido no ciclo de transferência da informação com o uso da informática se deu no plano da interoperabilidade, quando os registros documentais passaram a ser legíveis por máquina, gerando uma fonte de intercâmbio entre instituições que poderiam diminuir seus custos de processamento em suas bases de dados. No entanto, o volume de informações gerado pelo crescimento e dispersão da literatura técnica e científica em várias mídias mantinha o processo de comunicação lento, caro e cumulativo.

O cenário para o sistema de comunicação sem papel (Paperless Communication System) idealizado por Lancaster para o ano 2000, previa que cada produtor de informação teria um terminal *on-line*

em seu escritório e outro em sua residência, dotados de monitor e teclado, entre outros periféricos, para a transmissão da informação. Algumas formas de comunicação usariam também canetas óticas e toque com os dedos na tela. Os cientistas usariam o "terminal" para criar, transmitir e receber informação. Os suportes em papel desapareceriam, restando apenas aqueles que, intencionalmente, desejassem manter fisicamente os objetos por interesse pessoal ou social. Todo o processo comunicacional seria eletrônico, pessoal, ficando apenas sob responsabilidade de grandes empresas a tarefa de organizar as informações, quer dizer, gerenciá-las para a distribuição segundo os perfis pessoais dos usuários da grande rede de informações.

O modelo abaixo, concebido por Lancaster em 1977, ilustra os quatro níveis de processamento da informação: 1. Gerência e monitoramento, que controla toda a operação; 2. Acesso ao documento (texto completo); 3. Usuário institucional; e 4. Usuário pessoal.

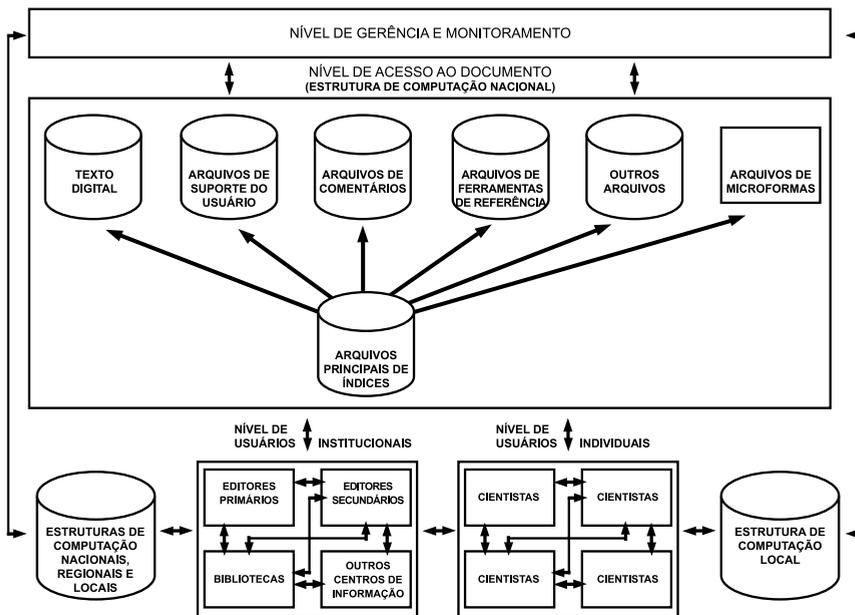



Figura 2 – Nível macro de configuração de um Sistema Nacional Digital para a comunicação tecno-científica
Fonte: Adaptada de Lancaster (1977, p. 18).

É evidente, pelo desenho do sistema, que a estrutura em níveis de processamento da informação considera a complexidade do envolvimento e da responsabilidade do indivíduo e do Estado, passando pelas instituições. No entanto, o modelo está longe de representar o cenário atual, da primeira década do século XXI, embora em suas considerações finais, Lancaster tenha listado 17 atividades que justificariam o surgimento do sistema de informação ou sociedade sem papel, principalmente a quarta que aponta para a criação da internet e da web.

> [...] 4. A emergência de computadores conectados em rede e de um trabalho constante de desenvolvimento de melhoria da capacidade das telecomunicações por satélite e outros meios. O desenvolvimento da transmissão digital é particularmente importante para a implementação de um sistema mundial de comunicação científica. (LANCASTER, 1977, p. 22, tradução nossa)[11]

Lancaster conclui seu artigo, sugerindo que essa mudança social nos processos de transferência da informação, do mundo analógico para o mundo virtual engendra questões de ordem tecnológica, intelectual, social e psicológica que podem ser representadas por preocupações de preservação da memória, gerenciamento da rede, diversidade de línguas e de linguagens, propriedade intelectual e mudança de atuação das instituições, principalmente das bibliotecas.

Em sua reflexão prospectiva, Lancaster apresenta questões que já foram resolvidas pelo desenvolvimento das tecnologias da informação e comunicação, como a criação da internet e da web, os equipamentos que deixam de ser terminais e passam por uma autonomia cada vez maior de interação informativa, a descentralização do controle de arquivos digitais e outras tantas questões operacionais. Em contraponto aos avanços tecnológicos após mais de 30 anos de suas previsões, ele nos deixa observações que ainda são consideradas o cerne do futuro da organização do conhecimento e das configurações da memória no ciberespaço.

---

11 “[...]4. The emergency of networks connecting computers, and continued work on the development of improved telecommunications capabilities by satellite and other means. Improved digital transmission capabilities are particularly important to the implementation of a worldwide system for science communication.”

A primeira delas diz respeito à possibilidade de que a transferência da informação se faça por canais informais, o que nos leva ao mundo da oralidade e da construção da memória coletiva operacionalizada pela narrativa oral, de caráter transformador e sem a preocupação de acumulação material ou "restos". Ainda neste aspecto, a representação do fluxo de informações adota o modelo de ciclo, o que reforça a idéia de transformação em contraste com a acumulação de informações. Em outro plano de observação, o gerenciamento da rede mundial de computadores ainda permanece sob discussão, assim como as questões de direitos autorais e de preservação das memórias acumuladas pela tecnologia da escrita.

Passados mais de 40 anos da construção deste modelo que influenciou o ensino e a pesquisa no campo da Ciência da Informação, procuramos na internet alguma referência àquele proposto por Lancaster, mas que indicasse modificações que ocorreram nas operações pertencentes à dinâmica da informação técnico-científica no mundo contemporâneo. Com a ajuda dos alunos do curso de Organização de conceitos em linguagens documentárias, da Universidade Federal do Estado do Rio de Janeiro (UNIRIO), elegemos o modelo *The (changed) Information Cycle* apresentado no Blog de George Siemens, com o título de *Elearnspasce: learning, networks, knowledge, technology, community* (SIEMENS, 2009).

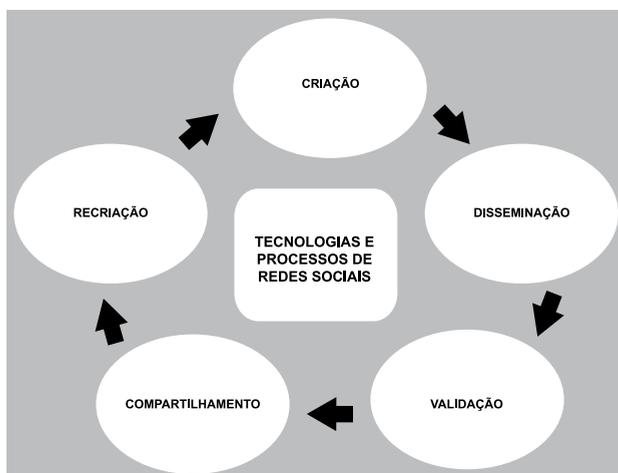



Figura 3 - Ciclo da Informação (modificado)
Fonte: Siemens (2009).

O *Blog* de Siemens é um espaço de aprendizagem eletrônica, o que explica a apresentação do modelo gráfico – Ciclo da informação (modificado). Resumidamente, o autor afirma que a informação é criada como produto da pesquisa, depois ela é disseminada através de conferências, discussões com colegas, etc. Em seguida, a validação e a exatidão da informação são atribuídas por pares especialistas, e se a informação for classificada como relevante, ela é então publicada em periódicos, conferências, resenhas e livros. E, ainda, se a informação/pesquisa se mantiver atual, ela é reproduzida em publicações posteriores e usada como base para o desenvolvimento daquele campo de conhecimento.

O autor comenta, ainda, que, surpreendentemente, os elementos do ciclo da informação não mudaram substancialmente em razão dos avanços tecnológicos. O processo, entretanto, tem sido enormemente alterado. Segundo Siemens, a <u>criação</u> é, hoje em dia, simplesmente representada pelos podcast ou postagens de blogs. A <u>disseminação</u> e a <u>avaliação</u> ocorrem pelos métodos de consulta aberta *on-line* (*crowd sourcing*), tais como escavação de links, índices no sítio Amazon, comentários no DIIGO[12] – que é uma ferramenta de pesquisa e uma comunidade de compartilhamento de conhecimentos –, entre outros. O <u>compartilhamento</u> dos recursos ocorre por meio de artigos online em periódicos de acesso livre, *blogs*, vídeos, postagens, construções no Second Life. A <u>recriação</u> (assimilação), quando os recursos são licenciados para permiti-la, ocorrem por meio de reformatação de conteúdo (mushups)[13] e transposição de formatos (repurposing content)[14] em várias mídias e línguas. Finaliza o autor, considerando que a informação é agora mutável, participativa, democrática e rapidamente recriada.

12 Ver http://www.diigo.com/

13 Ver http://blogs.microsoft.co.il/blogs/gadshor/archive/2008/03/19/mushup-what-s-the-origin-of-the-word.aspx

14 That means taking intellectual property you have created and using it in another way. For example, turning a book into an audio program, or excerpting parts of the book into articles. Repurposing content can help you reach a brand new audience with your ideas. One advantage of transferring your existing information into new formats is that you can meet the needs of people with different learning styles. This can help you to reach a larger audience with your message, as well as increasing your profits. Ver http://ezinearticles.com/?Repurposing-Content&id=110568

POR UMA MEMÓRIA CRIATIVA NO CIBERESPAÇO

Qual seria o papel dos Repositórios Institucionais e das Bibliotecas Digitais nessa dinâmica comtemporânea da transferência da Informação? Que parte da produção científica e tecnológica poderia deixar de ser preservada institucionalmente? Quem selecionará as informações que deverão fazer parte da memória científica no ciberespaço? Parece que as instituições devem começar a desenvolver, com urgência, políticas de seleção das informações que farão parte de seus repositórios institucionais e biblotecas digitais. Pensando bem, este tem sido o desafio já há longos anos das bibliotecas universitárias e mais recentemente dos arquivos das instituições de ensino.

Não temos resposta a todas essas questões, mas podemos sugerir alguns caminhos com base no que escolhemos para alimentar a discussão sobre a memória e o conhecimento. O universo da memória e sua interação com o mundo da informação; as considerações sobre as condições de produção do conhecimento no ciberespaço; e, as análise sobre o ciclo da informação em uma sociedade sem papel feita por Lancaster nos levam a apresentar três modelos de fluxos de informação que explicam formas possíveis de produção do conhecimento, na sociedade contemporânea.

Considerando a tensão entre as propriedades oralidade/visualidade dos recursos ou inscrições que transitam no espaço da memória e da informação; e, a tensão entre as operações acumulação /transformação da dinâmica da produção de conhecimento, se junta uma outra propriedade ao processo de transferência da informação, a reprodutibilidade que justifica, parcialmente, a formação de acúmulos, "restos" ou documentos.

1. Produção de conhecimentos – Assimilação ou re-criação (Oralidade)
2. Produção de conhecimentos – Registro – Assimilação ou re-criaçã (Visualidade)
3. Produção de conhecimentos – Registro – Documento – Assimilação ou re-criação (Reprodutibilidade)

A classificação é preliminar e apresenta, como em Lancaster, cenários de trânsito da informação com o objetivo de incentivar a produção de novos conhecimentos. O primeiro fluxo, característico da oralidade, não comporta registros e, portanto não alimenta a acumulação em suportes material ou digital. A transmissão da informação se dá apenas pela fala e, na medida em que é assimilada pelo receptor, imediatamente já é transformada em novos conhecimentos, na ordem do individual. Consideramos que este fluxo de informação se dá no espaço da virtualidade e como comenta Le Goff, nas sociedade sem escrita, a memória transmitida não é uma memória palavra por palavra. Goody (1977 apud LE GOFF, 2003, p. 426) afirma: "o produto de uma rememoração exata aparece nestas sociedades como menos útil, menos apreciável que o fruto de uma apreciação inexata." A memória coletiva parece funcionar nestas sociedades segundo uma reconstrução generativa e não como uma memorização mecânica, proporcionando, dessa forma, uma memória mais criativa.

Vale observar, no entanto, que caminhamos para uma oralidade tambem digitalizada, pois a comunicação entre as pessoas tem feito uso da convergência tecnológica em que usamos uma potente midia, a internet, que reúne televisão, radio, vídeo, além dos textos. Grande parte das trocas conversacionais são mediadas por tecnologias, o que ocasionaria uma fala digital, sujeita a restos.

O segundo fluxo possibilita o aparecimento de restos ou memórias auxiliares, mesmo que em suportes apenas digitais e sem garantias, por enquanto, de salvaguarda por um período longo. A visualidade é compreendida como a propriedade de dar a conhecer por meio visual e aí estariam incluídas as tecnologias da escrita e da imagem. Nunca se escreveu tanto como agora; uma escrita digital, com transmissão em tempo real, pronta para ser assimilada e transformada. A dinâmica da transferência da informação na esfera da visualidade digital, embora se aproxime da oralidade pela velocidade de transmissão, dela se distancia por oferecer ainda a possibilidade de reprodução. È nessa tensão entre história e memória que Jacques Le Goff nos diz que "o estudo da memória social é um dos meios fundamentais de abordar os problemas do tempo e da história, relativamente aos quais

a memória está ora em retraimento[dando margem ao acúmulo], ora em transbordamento [proporcionando a transformação]"(LE GOFF, 2003, p. 422).

O terceiro fluxo apresenta um custo muito elevado de manutenção e, portanto, tem fraca possibilidade de existência futura no ciberespaço. Nele estão contidos os processos de seleção, aquisição, manutenção e substituição. Esses restos, ou acúmulo de registros representam hoje as memórias auxiliares com as quais nos habituamos a conviver desde o aparecimento da escrita e da história, momento em que adicionamos à oralidade os registros do conhecimento e sua manutenção ou proteção patrimonial.

As instituições devem representar a garantia da qualidade das informações colocadas em seus seus repositórios e bibliotecas digitais. A seletividade, nesses caso, é imperiosa para proteger documentos considerados relevantes para a sociedade por um período mais longo de tempo. O custo de manutenção das informações é elevadíssimo se considerarmos a substituição de midias, indexação e recuperação de dados. Não basta preservar apenas os endereços, como é caso do Internet Archives, mas é necessário poder encontrar a informação por outros indicadores como autoria, assunto, conteúdo enfim.

CONSIDERAÇÕES FINAIS

Há ainda muitos desafios para a organização do conhecimento e muito a pesquisar sobre o tema, considerando a velocidade com que mudam as formas e os meios de comunicação na sociedade. Dos quatro problemas apontados por Lancaster no ano de 1977 em relação à eficiência da transferência da informação: 1. a literatura que cresce a taxas exponenciais; 2. a dispersão dessa literatura em mídias diversas; 3. o espaço de tempo cada vez maior entre o resultado da pesquisa e sua divulgação; e, 4. os custos de publicação que ultrapassam os orçamentos do pesquisador individual e das instituições de guarda e disseminação da informação, podemos dizer que os dois primeiros se mantêm (não mais como um problema), tal é a quantidade de

informação lançada na rede mundial. Os dois últimos, ao contrário, vêm sendo solucionados pelas facilidades oferecidas de publicação e acesso livre à informação.

Considerando como um novo paradigma na organização do conhecimento, a facilidade no acesso e disseminação da informação, aliada à proteção dos itens das coleções fazem dos repositórios institucionais modelos híbridos de patrimônio da humanidade que não disputam espaço com outros arquivos e bibliotecas digitais, apenas somam em qualidade de oferta.

Como a criança que experimenta o mundo, provando um pedaço de todas as coisas para armazená-las na memória, o Patrimônio da Humanidade passa à adolescência e reconhece que a acumulação desenfreada de traços de memória poderá levar a sociedade a uma incapacidade de recuperar ou de atualizar os acontecimentos do passado. Ao lado do necessário movimento "disseminador" de lançar, a cada fração de segundo, objetos na rede mundial de computadores surgem ações patrimoniais digitais "protetoras", que estão preocupadas com a qualidade desses objetos, como exemplo, os repositórios institucionais e a Biblioteca Digital Mundial. Ainda que essas ações estejam ancoradas no espírito da disputa por visibilidade e recursos ou de memória do mundo em seu sentimento democrático, elas são instituídas formalmente e significam escolhas que revelam a prudência de qualificar os objetos que deverão ser protegidos na memória no ciberespaço. Inaugura-se, assim, uma linha de fuga, ou um espaço criativo de memória, com a compreensão do equilíbrio entre lembrar e esquecer.

REFERÊNCIAS

BIBLIOTECA DIGITAL MUNDIAL. Disponível em: <http://www.wdl.org/pt/> Acesso em: 23 de jan. 2010.

BILLIGNTON, James H. La Bibliothèque numérique mondiale réduit la fracture culturelle. **Le Monde**, Paris, 24 avr. 2009.

BUCAILLE, Richard; PESEZ, Jean-Marie. Cultura material. In: **Enciclopédia**

Einaudi, Lisboa: Imprensa Nacional; Casa da Moeda, 1989. p. 11-47. v. 16.

DAVIS, Philip M.; CONNOLLY, Mathew J. L. Institutional repositories: evaluating the reasons for no-use of Cornell University's installation of DSpace. **D-Lib Magazine**, v. 13, n. 3/4, 2007.

DIGITAL ACCESS TO SCHORLARSHIP AT HARVARD. Disponível em: <http://dash.harvard.edu/handle/1/1> Acesso em: 23 de jan. 2010.

DODEBEI, Vera. Espaço mítico e imagético da memória social. In: COSTA, Icléia Thiesen Magalhães; GONDAR, Jô (Org.). **Memória e espaço**. Rio de Janeiro: 7 Letras, 2000. p. 63-71.

______. Memória e conhecimento: oralidade, visualidade e reprodutibilidade no fluxo da informação. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO - ENANCIB, 10., João Pessoa. **Anais ...**, João Pessoa: ANCIB, 2009. 1 CD-ROM.

DSPACE. In: WIKIPEDIA. Disponível em: <http://en.wikipedia.org/wiki/DSpace> Acesso em 21 de jan. 2010.

EUROPEANA. Disponível em: <http://www.europeana.eu/portal/aboutus.html.> Acesso em: 23 de jan. 2010.

GONDAR, Jô. Entre lembrar e esquecer: desejo de memória. In: COSTA, Icléia Thiesen Magalhães, GONDAR, Jô. (Org.). **Memória e Espaço.** Rio de Janeiro: 7 Letras, 2000. p. 35-43.

GOOGLE BOOKS. Disponível em: <http://books.google.com/> Acesso em: 23 de jan. 2010.

HALBWACHS, Maurice. **Memória coletiva**. São Paulo: Centauro, 2004.

INTERNET ARCHIVES. Disponível em: <http://www.archive.org/> Acesso em 23 de jan. 2010.

LANCASTER. Frederic Wilfrid. **The dissemination of scientific and technical information**: toward a paperless system. Illinois: University of Illinois at Urbana-Champaign, 1977. Occasional Papers, n. 127. University of Illinois at Urbana-Champaign Library. Large scale Digitization Project, 2007.

LE GOFF, Jacques. **História e memória**. Campinas: Editora da Unicamp, 2003.

LEVIE, Françoise. **L´homme qui voulait classer le monde**: Paul Otlet et le Mundaneum. Bruxelles: Impressions Nouvelles, 2006.

LÉVY, Pierre. **A inteligência coletiva**: por uma antropologia do ciberespaço. São Paulo: Loyola, 2003.

NORA, Pierre. Entre memória e história: a problemática dos lugares. **Projeto História**, Revista do programa de estudos pós-graduados em História e do Departamento de História. São Paulo, n. 10, p. 1-78, dez. 1993.

OPEN DOORS AND OPEN MINDS: what faculty authors can do to ensure open access to their work through their institution. Washington, DC, SPARC, 2008.

(A SPARC / SCIENCE COMMONS WHITE PAPER, April 2008). Disponível em: <http://www.arl.org/sparc/bm~doc/opendoors_v1.pdf> Acesso em: 22 de jan. 2010.

PIEDADE, Maia Antonieta Requião. **Introdução à teoria da classificação**. 2. ed. rev. e aum. Rio de Janeiro: Interciência, 1983.

ROGET, Peter Mark. **Roget's Thesaurus of English words and phrases**. London: Peguin Books, 1998. [1. ed. 1852]

SETTIS, Salvatori. Warburg continuatus: descrição de uma biblioteca. In: BATATAIN, Marc; JACOB, Christian (Org.). **O poder das bibliotecas**: a memória do livro no ocidente. Rio de Janeiro: Editora da UFRJ, 2000. p. 108-154.

SIEMENS, George. **Elearnspace**. Disponível em: <http://www.elearnspace.org/blog/2009/04/17/the-changed-information-cycle/> Acesso em: 23 de jan. de 2010.

WARE, Mark. Institutional repositories and scholarly publishing. **Learned Publishing**, v. 17, n. 2, 2004, p. 115-124. Disponível em: <http://www.alpsp.org/ngen_public/> Acesso em 22 de jan. 2010.

YATES, Francês A. **A arte da memória**. Tradução de Flávia Bancher. Campinas: Editora da Unicamp. 2007.

A experiência da UFPR na construção de repositórios digitais:
a implantação integrada das ferramentas Dspace e Open Journal System.

Marcos Sunye
Ligia Setenareski
Fabiano Silva
Edson Ramiro
Lucio Foltran

INTRODUÇÃO

A criação do ARXiv[1] é vista como o primeiro exemplo de concepção de um arquivo digital global para arquivos acadêmicos. Criado inicialmente para a área de Física, logo incorporou também Matemática e Ciência da Computação e funciona hoje na Universidade de Cornell. Os princípios definidos na criação do ArXiv no início dos anos 1990, foram incorporados pela Open Archives Initiative (OAI)[2] que busca facilitar a publicação da produção científica e fornecer o acesso geral às comunidades interessadas nessas publicações. Desde as primeiras reuniões ao final dos anos 1990 até os dias de hoje a OAI ganhou apoio da maioria das universidades e centros de pesquisa além de ter seu protocolo de comunicação OAI-PMH (Protocolo for Metadata Harvesting) incorporado à maioria dos *softwares* de gestão de bibliotecas digitais.

A implantação da Biblioteca Digital da Universidade Federal do Paraná (UFPR) seguiu, desde o seu início em 2004, os mesmos objetivos

1 http://arxiv.org/

2 http://www.openarchives.org

da OAI, facilitando a publicação para a sua comunidade científica e tornando universal o acesso à estas publicações. O acervo inicial era composto por teses e dissertações (TeDs) e a partir de 2005 a UFPRTV passou a fazer parte da Biblioteca Digital disponibilizando sua programação em vídeos. Nesse mesmo ano foi implantado o repositório das revistas científicas produzidas pela UFPR, completando assim o acervo digital que permanece até os dias atuais.

A Biblioteca Digital da UFPR faz uso do protocolo OAI-PMH na disponibilização de seu acervo. Este protocolo é, por sua vez, utilizado tanto por grandes repositórios científicos, como o OAIster[3] quanto por mecanismos de busca, como o Google, garantindo assim a universalidade do acesso.

A UFPR usa dois *softwares* em sua biblioteca digital, o primeiro é o *software* Dspace[4] que, após um estudo comparativo com outros *softwares* de gerência de acervos digitais, foi escolhido por sua robustez e sua adequação às dimensões da UFPR que tem uma média de produção anual de 1000 teses. O Dspace é o *software* responsável pelas TeDs e os vídeos da TV UFPR. O segundo é o *software* Open Journal System (OJS)[5] que tornou-se uma referência na gerência de revistas científicas de acesso livre. O OJS é o *software* de gerência das 33 revistas científicas da UFPR.

A UFPR desenvolveu seu acervo digital integrado ao acervo convencional do Sistema de Bibliotecas (SiBi), o que permite uma busca única ao acervo físico e digital. Os metadados são gerados uma única vez no *software* de gerenciamento utilizado pelo SiBi e posteriormente são reproduzidos na biblioteca digital. Esta integração da biblioteca digital ao *software* de busca aos catálogos do acervo físico do SiBi ocorreu já no processo de implantação do *software* Virtua, iniciado em 2001. Em 2009 houve a migração do *software* Virtua para o *software* Sophia, com a manutenção da integração dos bancos de dados.

Atualmente a biblioteca digital da UFPR é responsável por boa parte do conteúdo acessado dentro do domínio ufpr.br e já ultrapassou

---

3 www.oaister.org

4 http://www.dspace.org

5 http://pkp.sfu.ca/?q=ojs

os 10.000 itens. Este texto descreve a estrutura geral da biblioteca digital, as soluções adotadas e o impacto das publicações digitais na comunidade acadêmica.

INFRAESTRUTURA UTILIZADA PARA PROCESSAMENTO, ARMAZENAMENTO E RECUPERAÇÃO DO ACERVO DIGITAL DA UFPR

As aplicações que compõem o Acervo Digital da UFPR encontram-se distribuídas em dois servidores distintos e fisicamente separados. O servidor chamado de primário (Figura 1) agrupa os *softwares* Dspace, Open Journal System e Open Conference System com suas respectivas bases de dados. O servidor chamado de secundário hospeda o *software* Sophia de gerência do acervo físico e de empréstimos, além de armazenar os metadados das TeDs, das revistas e dos vídeos. O secundário é o que armazena todos os objetos digitais assim como a réplica dos respectivos metadados.

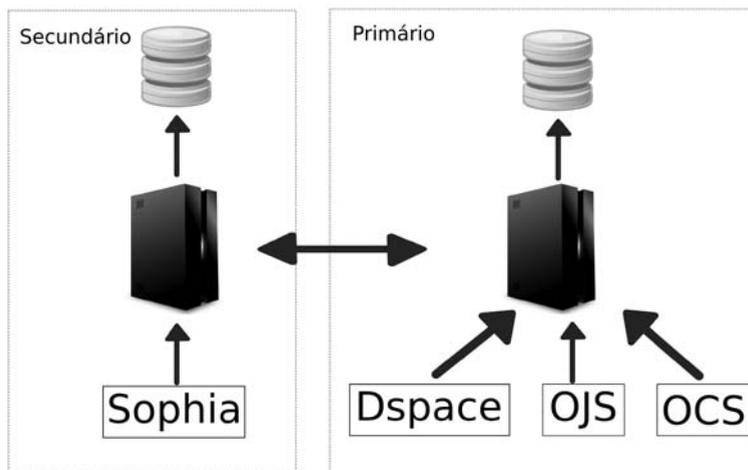



Figura 1 - Distribuição dos servidores de suporte ao Acervo Digital da UFPR.

O servidor primário é um Opteron com 32Gb de memória RAM e 9Tb de disco. O ambiente adequado para o funcionamento desta solu-

ção inclui um servidor de páginas Web, um servidor de e-mail e dois sistemas gerenciadores de banco de dados (SGBD). Como servidor de páginas é utilizado o Apache Http Server[6] e como SGBDs são utilizados o PostgreSQL[7] e o MySQL[8]. Todos estes *softwares* rodam sobre o sistema operacional Debian GNU/Linux[9]. Todo o ambiente computacional é composto por sistemas de fácil atualização, com suporte contínuo e de qualidade.

DSPACE

O Dspace é um *software* que foi inicialmente desenvolvido por um grupo do Massachusetts Institute of Technology (MIT) e da Hewlett-Packard Labs na Universidade de Cambridge. Atualmente seu desenvolvimento é feito voluntariamente por meio de uma comunidade de usuários, desenvolvedores e pesquisadores. Desde sua criação, o foco principal é a manutenção e gerência de repositórios digitais, sejam estes particulares ou institucionais.

O primeiro critério utilizado na escolha do Dspace para gerenciar a biblioteca digital das TeDs e dos vídeos produzidos pela UFPR foi o seu desempenho. Entendido por desempenho a sua capacidade de armazenar objetos digitais sem limites de número de objetos nem tamanho dos mesmos. Em seguida a possibilidade de se criar coleções independentes com a possibilidade de se adequar cada uma delas aos padrões de armazenamento, preparo e recuperação da informação adotados pelo SIBI.

O Dspace é um *software* livre registrado sob licença General Public License (GPL)[10], o que é um grande facilitador que elimina a necessidade de contratos de aquisição e manutenção. Além disso, sua comunidade é a mais ativa entre os gerenciadores de acervo digital o que proporciona agilidade, resolução dos problemas e boa qualidade de

6 http://httpd.apache.org

7 http://www.postgresql.org

8 http://www.mysql.org

9 http://www.debian.org

10 http://www.gnu.org/licenses/gpl.html

documentação sobre o uso deste *software*. Em meia década de utilização a UFPR não teve problemas de perda de acervo, corrupção de arquivos ou falhas de segurança em seu repositório Dspace que mostrou ser um sistema adequado às necessidades da UFPR.

OPEN JOURNAL SYSTEM

Essa ferramenta propõe o gerenciamento completo de toda a produção de uma revista eletrônica. É um sistema de código aberto concebido pelo Public Knowledged Project, criado pela parceria entre a Simon Fraser University, Stanford University e a Arizona State University e distribuído sob a licença GPL.

O fato de ser um *software* livre possibilitou a tradução e diversas melhorias, feitas tanto pela equipe do Acervo Digital da UFPR, como pela equipe do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT).

O OJS foi implantado na UFPR para gerenciar a editoração das revistas científicas produzidas. Antes da implantação do OJS na UFPR, algumas revistas faziam submissões de artigos de uma maneira totalmente manual, isto é, enviando o artigo impresso para o avaliador, que por sua vez corrigia tal versão, reenviando-a para o editor. Tal processo poderia ocorrer inúmeras vezes até que todas as correções exigidas fossem feitas ou a não aceitação do artigo fosse explicitada. Esse processo manual tornava inviável a produção de revistas em curto prazo de tempo.

Com a informatização do processo, a editoração das revistas ganhou mais velocidade e flexibilidade, pois toda a comunicação entre as partes passou a ser feita por e-mail e como consequência positiva o processo de avaliação dos artigos ficou mais confiável. Outra característica importante é a possibilidade de guardar um histórico dos e-mails e de todos os documentos que passam pelo sistema, tornando possível uma futura auditoria.

O Virtua[11] é um sistema proprietário de gerenciamento de bibliotecas, desenvolvido pela Visionary Technology in Library Solutions (VTLS). De acordo com a VTLS, o Virtua é flexível e capaz de atender necessidades individuais de cada instituição cliente.

Entre o início de 2001 e início de 2009 este *software* foi o responsável por administrar o acervo físico, e também por realizar um controle automatizado de empréstimos do acervo e dos usuários cadastrados no SiBi[12]. Estes foram os motivos que levaram a sua implantação em todas as bibliotecas do SiBi/UFPR, a universidade necessitava de uma ferramenta que fosse capaz de cumprir estas tarefas, e o Virtua as realizava de forma satisfatória.

Entretanto, o encerramento do contrato e o alto custo de manutenção do Virtua inviabilizaram a continuação de seu uso e fez com que a administração do SiBi buscasse uma solução mais viável economicamente, sem perder o atendimento aos requisitos mínimos pré-estabelecidos, e a melhor solução foi apresentada pelo *software* Sophia.

O Sophia[13] é um sistema proprietário de gerenciamento de bibliotecas, desenvolvido pela Primasoft. A implantação do Sophia foi feita de forma gradual. À medida em que os componentes eram instalados, os mesmos eram verificados para dar prosseguimento aos próximos passos como forma de garantir as funcionalidades do sistema. O último e mais importante passo desta implementação se deu quando todas as informações de livros, cadastros de usuários, entre outras, foram devidamente migradas do Virtua para o Sophia. Uma vez instalado e funcionando o Sophia, foi necessário fazer a integração com o Dspace.

11 http://www.vtls.com/products/virtua

12 http://www.portal.ufpr.br

13 http://www.primasoft.com.br/2006/html/interna_1.php?cod=29

INTEGRAÇÃO DO SISTEMA GERENCIADOR DE BIBLIOTECA (SGB) COM O SISTEMA GERENCIADOR DE REPOSITÓRIO DIGITAL (SGRD)

A integração SGB/SGRD na UFPR é motivada por três fatores fundamentais:

- a inexistência de pessoal dedicado ao SGB na UFPR – sendo assim a indexação das TeDs eletrônicas deveria ser automaticamente derivada da indexação das TeDs convencionais evitando-se assim o aumento dos recursos humanos envolvidos;
- ambos os acervos (SGB e SGRD) deveriam ser sincronizados, garantindo-se assim uma busca unificada;
- o padrão de indexação deveria ser o mesmo com o seu controle sendo centralizado no sistema de bibliotecas.

A primeira integração foi realizada ao antigo sistema usado, o Virtua, que não possuía facilidades para que este processo fosse executado. Cada nova tese ou dissertação disponível também em meio eletrônico era catalogada dentro do sistema Virtua como um documento convencional. O processo de integração, desenvolvido pela UFPR, constituía em realizar uma conexão com o servidor que hospedava o antigo sistema e utilizar o protocolo OAI-PMH para realizar um *harvest* em busca de atualizações recentes. O *harvest*, que é um processo de busca de metadados por meio da utilização de um protocolo próprio, coletava todos os metadados alterados no último dia e estes eram colocados um a um em diretórios separados. Ao mesmo tempo em que os metadados eram divididos, eram convertidos de Machine-Readable Cataloguing (MARC)14 para o padrão Dublin Core15. Todo este processo era realizado utilizando bibliotecas Java para *parser* de arquivos XML. Finalmente separados e organizados, os itens eram adicionados por uma classe de *import* do Dspace e poderiam ser acessados pela página Web. Porém, neste ponto, só estavam disponíveis os metadados

14 http://www.loc.gov/marc

15 http://dublincore.org

da tese/arquivo. A inserção do arquivo em formato pdf, denominado pelo Dspace de *bitstream* deveria ser feito de forma manual por uma pessoa que possuísse uma conta no Dspace e fosse administrador da coleção a qual o item e seu *bitstream* deveria ser inserido.

Esta submissão assíncrona entre os metadados e o arquivo digital em formato pdf foi resolvida na troca do SGB, com a substituição do *software* Virtua pelo *software* Sophia e uma extensão do Sophia desenvolvida especialmente para assegurar o processo de integração.

Atualmente é necessário apenas uma única intervenção para que a tese/arquivo seja disponibilizada tanto no SGB quanto no SGRD. Quando uma alteração é realizada na interface de administração do Sophia, este armazena uma cópia dos metadados e realiza um contato com uma aplicação PHP que roda no mesmo servidor que o Dspace para que o *bitstream* e os metadados sejam adicionados ao mesmo tempo no Dspace. Com o sucesso da operação, a aplicação devolve para o Sophia o respectivo *link* de acesso ao item, o que possibilita a busca por teses/arquivos diretamente no SGB.

Devido ao fato do SGB já possuir uma extensão exclusiva para integração com o Dspace não foram detectados grandes problemas, as alterações hoje são realizadas em tempo real e não mais diariamente como era feito no Virtua. Outro fator importante desta integração é que não há mais a necessidade de uma posterior inserção do *bitstream*, isto já é feito no momento da inserção dos metadados.

FLUXO DA INTEGRAÇÃO DO ACERVO DE TEDS

O fluxo de integração das TeDs assegura a sincronia entre o acervo convencional (Sophia) e o acervo de teses eletrônicas (Dspace). Ao realizar uma inserção no Sophia, os metadados são armazenados no servidor secundário e uma cópia destes é transferida para o servidor primário por meio de uma aplicação Web. Realizada a transferência, os dados são adaptados ao formato usado pelo Dspace e então a inserção é realizada. Caso haja sucesso nesta operação, a aplicação devolve ao Sophia o respectivo endereço de acesso ao item para que, à partir do Sophia, seja possível obter acesso ao item armazenado no Dspace de

forma transparente. Toda e qualquer tipo de alteração é feita a partir do próprio Sophia, que com o uso da aplicação web, faz a transferência e execução das alterações no servidor Dspace. Estas seguem exatamente o mesmo fluxo de execução de uma inserção, mudando apenas a operação a ser executada.

FLUXO DA INTEGRAÇÃO DO ACERVO DE VÍDEOS

O fluxo de integração do acervo de vídeos assegura a transferência dos vídeos da base de dados da TV até uma base de dados temporária no servidor primário para sua posterior indexação e inserção no Dspace. Uma aplicação web, desenvolvida pela UFPR permite que o bibliotecário responsável pela indexação possa ter acesso aos vídeos. A indexação é feita diretamente no Dspace e assim que os metadados são gerados, a pessoa pode realizar a submissão do vídeo no repositório. Esta submissão consiste em recuperar o vídeo as quais estas se referem, formatar e finalmente inserir no Dspace. A aplicação web desenvolvida tem por objetivo facilitar o processo de geração dos metadados e inserção de vídeos dentro do repositório. Para realizar alterações, modificações e remoções é necessário que o usuário possua uma conta no Dspace com privilégios de administrador da coleção e siga as orientações do próprio Dspace para a realização destas tarefas.

FLUXO DE INTEGRAÇÃO DO ACERVO DE REVISTAS CIENTÍFICAS

O processo de submissão de um artigo para publicação em alguma das revistas hospedadas na biblioteca digital segue o roteiro padrão do Open Journal System. O usuário se cadastra no sistema como um autor na revista em que deseja publicar e, após o cadastro feito, tal usuário entra no sistema e faz a submissão do artigo.

No momento da submissão do artigo o autor preenche um formulário com informações sobre o artigo e estas informações são usadas para gerar os metadados deste artigo. Após a geração dos metadados é possível enviar o artigo para avaliação.

O editor da revista recebe o artigo e designa-o para a revisão da

submissão, fazendo o acompanhamento do artigo em todo o processo dentro do sistema, ou designa-o a um editor de sessão que fará este acompanhamento.

Antes de o editor encaminhar o artigo, ele verifica se todos os dados pedidos pela revista foram preenchidos e caso não tenham sido preenchidos, ele pede ao autor que complemente-os. Se os dados estiverem corretamente preenchidos, o artigo será encaminhado para o processo de revisão da submissão.

É na revisão da submissão que o editor ou editor de sessão convida revisores para comentarem e revisarem o artigo. Esta rodada pode ser feita inúmeras vezes até que o artigo satisfaça todos os pedidos feitos pelos revisores. Nesta etapa o autor pode fazer modificações e reenviá-las para os revisores carregando um novo arquivo no sistema. Caso o artigo seja aceito, ele passa para o editorial.

O editorial, dependendo da revista, pode ser uma pessoa, uma equipe, ou uma empresa terceirizada. Nesta etapa é feita a diagramação e o leiaute e, após todas estas etapas, o editor da revista agenda em qual edição o artigo será publicado.

Na UFPR as publicações das edições são feitas em formato pdf. O OJS suporta outros formatos para publicação além do pdf, como o HTML, PS e outros. A publicação de uma edição a torna disponível na página da própria revista.

A subsequente disponibilização dos metadados dos artigos através do protocolo OAI-PMH assegura a visibilidade do acervo de revistas da UFPR. Os metadados ainda não são replicados automaticamente no Sophia já que são gerados pelos próprios autores e estão fora do padrão utilizado pelo sistema de bibliotecas. A sua inserção ainda depende de uma revisão por parte de um bibliotecário garantindo assim a qualidade dos metadados.

Uma primeira análise Ruthes (2007) demonstrou que a criação dos metadados pelos autores gera uma série de problemas e entre eles a falta de precisão nas buscas. Por esta razão a UFPR tem aprofundado os estudos de melhoria da qualidade dos metadados gerados pelos autores antes de efetivar a integração automática entre o acervo digital e o acervo convencional de revistas (MARTINHAGO, 2006).

DISPONIBILIDADE E CONFIABILIDADE DO ACERVO DIGITAL

A manutenção de um acervo digital permanente com alta confiabilidade e disponibilidade pode implicar em uma progressão geométrica dos custos. A biblioteca digital ocupa hoje 7 *Terabytes* em disco e soluções corporativas de armazenamento deste porte ultrapassam 500.000 reais. A UFPR optou por uma hierarquia de servidores e a adoção de uma política de confiabilidade baseada na redundância de cópias.

Diariamente é efetuada uma cópia do banco de dados de cada aplicação no próprio disco do servidor que hospeda o acervo digital. Além desta cópia é feita uma outra, também diária, de todos os dados armazenados no SGBD. Com estas cópias pode-se garantir que caso ocorra alguma falha no SGBD, a recuperação possa ser feita de forma rápida e eficiente assegurando a disponibilidade do acervo digital.

Além deste procedimento de redundância local, para a recuperação rápida do servidor em caso de falha a cada 3 dias é executada uma cópia incremental não só dos arquivos armazenados, mas também das aplicações em um servidor localizado em um espaço físico diferente. A cada semana uma segunda cópia incremental copia os dados em um terceiro servidor. A redundância de servidores ao invés de uma única máquina de alta disponibilidade tem se mostrado eficiente e de baixo custo.

A INFLUÊNCIA DO REPOSITÓRIO DIGITAL NA VISIBILIDADE INSTITUCIONAL

Recentemente, metodologias de avaliação do impacto da produção científica têm levado em conta o conteúdo digital disponibilizado pelas universidades na internet como critério de avaliação. O Ministério da Educação da Espanha, através do Consejo Superior de Investigaciones Científicas (CSIC) elaborou uma metodologia de avaliação da visibilidade do acervo digital das instituições acadêmicas que gera um *ranking* com universidades de diversos países e avalia a quantidade de

publicações feitas por cada uma destas na Internet, o Webometrics (CSIC, 2009). A UFPR participa deste *ranking* porque seu acervo digital utiliza alguns dos protocolos de indexação abertos, como o Google Scholar[16] e o OAI-PMH, para a disseminação de suas produções.

Para fazer tal *ranking* são usados quatro indicadores: Tamanho, Visibilidade, Relevância e Google Scholar. O Tamanho é o número de páginas obtidas pelas ferramentas de busca como, Google Scholar, Yahoo, Live Search e Exalead; a Visibilidade é número total de *links* externos únicos citando uma página, que é obtido pelo Yahoo Search; a Relevância é a quantidade de arquivos publicados em formatos padrão (pdf, ps, doc, ppt); e o indicador Google Scholar é o número de artigos e citações para cada domínio acadêmico feitos pelo Google Scholar.

Cada índice é um *ranking*, isto é, para cada índice é feito uma tabela de posições entre as universidades participantes para que, posteriormente, seja usada a posição de cada instituição no cálculo da posição geral.

Os respectivos pesos para cada índice são: Tamanho = 20%, Visibilidade = 50%, Relevância = 15%, Scholar = 15%, somando cada parcela temos a posição geral de uma dada universidade.

Tamanho (20%) + Visibilidade (50%) + Relevância (15%) + Scholar (15%)

No ano de 2009 a UFPR obteve os seguintes valores para cada índice:

- Tamanho: 610
- Visibilidade: 663
- Relevância: 487
- Google Scholar: 29

Com tais índices a UFPR ficou na 352ª colocação dentre as 6.000 universidades participantes.

O ACERVO DIGITAL DA UFPR EM NÚMEROS

Atualmente a universidade disponibiliza por meio do Dspace o acesso a 4800 teses em formato digital produzidas por sua comunidade aca-

16 http://scholar.google.com

dêmica e mais de 600 vídeos produzidos pela UFPRTV. Estes números estão em constante alteração, uma vez que diariamente o acervo é atualizado. Com relação a espaço físico de armazenamento os valores chegam próximos dos 3000 GB disponibilizados à comunidade. Além disto, ainda se encontram em processo de catalogação mais 4000 GB de vídeos.

Desde meados de fevereiro até meados de dezembro de 2009, o Dspace acumula mais de 72.000 buscas dentro do acervo, o que resulta em aproximadamente 240.000 acessos a itens (um item é uma página que contém todos os metadados e um *link* de acesso ao arquivo referido) e aproximadamente 523.000 visualizações de arquivos. A média é de 20.000 visualizações de arquivos/mês com picos em alguns meses de 40.000 visualizações de arquivos e 20.000 visualizações de itens.

Na figura abaixo é possível verificar a evolução da Biblioteca Digital em 2009:

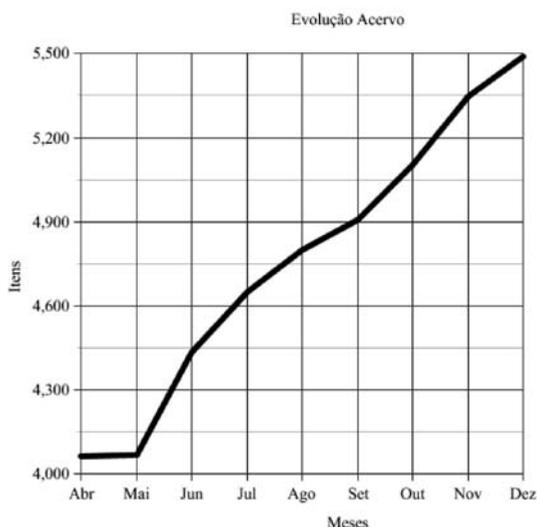



Figura 2 - Crescimento do número de itens disponibilizados pelo Dspace.

A página da biblioteca possui uma média de 1.800 visitas por dia e pouco mais de 51.000 visitas por mês. A distribuição de acessos de acordo com o país de origem pode ser vista no gráfico abaixo:

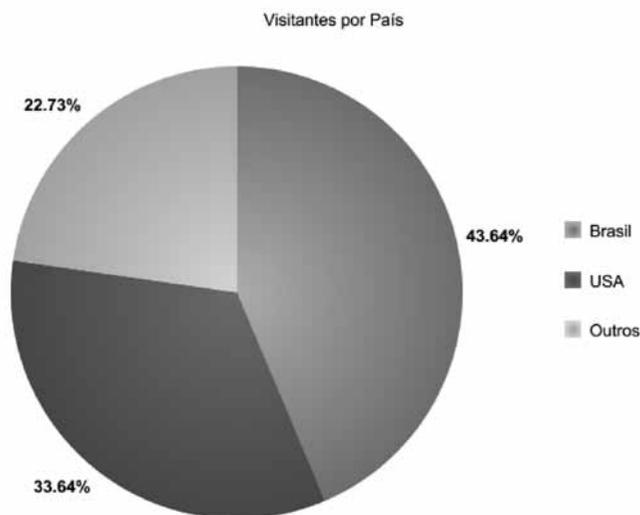



Figura 3 - Quantidade de acessos feitos ao Dspace por país de origem.

Revistas Eletrônicas

O acervo digital de revistas da UFPR é composto por 44 revistas, sendo 33 correntes. Os assuntos cobrem diversas áreas do conhecimento, tais como Biologia, Saúde, História, Filosofia entre outras. Algumas revistas publicam artigos em língua estrangeira e a mais usada delas é o inglês, seguido do francês, espanhol, italiano e alemão. A revista *Engenharia Térmica* é a única a ter publicações somente na língua inglesa.

Tais revistas contam com o apoio técnico do Centro de Computação Científica e Software Livre (C3SL)[17] e de um profissional especializado tanto no *software* OJS quanto no processo de editoração das revistas.

A periodicidade das edições das 33 revistas correntes varia entre bimestrais, trimestrais, semestrais e anuais. Deste total, 28 são também impressas e 5 mantém somente a publicação no formato eletrônico. No caso das revistas que são publicadas nos dois formatos é necessário um ISBN para cada versão.

A média de acesso a este acervo é de 9.132 acessos diários únicos e 269.789 visitas por mês. Grande parte dos acessos realizados de fora

17 http://www.c3sl.ufpr.br

do Brasil vem dos Estados Unidos, seguido de Portugal, Itália, França, México, Espanha, Alemanha, entre outros.

As revistas mais acessadas são a *Revista Brasileira de Geociências, Floresta* e a *Scientia Agraria* com média de 7.261 acessos por mês. Esta Biblioteca Digital dispõe de uma base de dados de 1Gb e 55Gb de arquivos em disco.

CONCLUSÃO E PERSPECTIVAS FUTURAS

A biblioteca digital da UFPR funciona hoje em um ambiente estável e dentro de um crescimento controlado. O uso amplo de *software* livre, desde o sistema operacional até a gerência do acervo proporciona uma flexibilidade gerencial visto que todo o sistema funciona sem necessidade de contrato de licença de uso.

O uso do ambiente formado pelos *software* Dspace e OJS superou as expectativas iniciais, considerando que além das vantagens do uso de software de código aberto, com a possibilidade de construção de extensões que facilitaram o processo de integração com o sistema de gerência do acervo convencional, contou-se com duas ferramentas estáveis e confiáveis. O dinamismo das comunidades Dspace e OJS são outro fator muito positivo que permite troca de soluções, documentação e experiências, além de uma rápida solução para eventuais *bugs*.

O maior risco à continuidade do acervo digital é a necessidade de manutenção de um quadro de pessoal altamente especializado, já que questões como segurança, confiabilidade e disponibilidade exigem procedimentos altamente complexos na área de computação.

A evolução do acervo digital passa agora por várias ações e entre elas podemos destacar:

- incremento da qualidade dos metadados gerados pelos autores e editores das revistas científicas até um padrão equivalente ao utilizado nas teses e dissertações, sem onerar os recursos humanos;
- implantação de um sistema que informa os autores dos artigos científicos sobre número e origem dos *downloads* de seus trabalhos.

- implantação de uma rede de *self archiving* para os pesquisadores da UFPR, para que possam arquivar e disponibilizar as publicações de sua autoria já publicadas em outros veículos;
- digitalização do acervo retroativo de teses e dissertações e de revistas.

REFERÊNCIAS

CSIC - CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS. **Ranking mundial de universidades em la red**. Jul. 2009. Disponível em: <http://www.webometrics.info/top6000_es.asp>. Acesso em: 12 ago. 2010.

MARTINHAGO, Adriana Zanhela. **Customização em ambientes de qualidade de dados**. 2006, Dissertação (Mestrado em Informática) - Universidade Federal do Paraná, Curitiba, 2006. Orientador: Marcos Sfair Sunye.

RUTHES, Eduardo Rocha. **Análise da qualidade de metadados em um sistema integrado de Bibliotecas Digitais**. 2007. Dissertação (Mestrado em Informática) - Universidade Federal do Paraná, Curitiba, 2007. Orientador: Marcos Sfair Sunye.

Dos processos analógicos às tecnologias digitais contemporâneas de recuperação da informação:

caminhos cognitivos na mediação para o acesso ao conhecimento

Aida Varela Varela[1]

Marilene Lobo Abreu Barbosa[2]

Igor Baraúna Guimarães[3]

INTRODUÇÃO

O termo repositório e os processos de organização da informação, temática na qual este artigo está inserido, têm sido recorrentes, nos últimos tempos, no âmbito da Ciência da Informação (CI), da Biblioteconomia e da Tecnologia da Informação (TI), isto porque as tecnologias digitais têm criado incessantemente soluções inovadoras, provocando mudanças constantes nas formas de organização da informação e na comunicação científica, acadêmica e da sociedade em geral. Daí, então, a sequência de aperfeiçoamentos que têm surgido nestas áreas, passando pelos bancos de dados eletrônicos, os Sistemas de Recuperação da Informação (SRI), as bibliotecas digitais e virtuais, os portais corporativos de informação etc.

---

1 Doutora em Ciência da Informação pela Universidade de Brasília (UnB). Docente Adjunto III do Instituto de Ciência da Informação. (varela@ufba.br)

2 Mestre em Ciência da Informação pelo Instituto de Ciência da Informação da Universidade Federal da Bahia. Professor do Instituto de Ciência da Informação da UFBA. (marilene@ufba.br)

3 Mestrando em Ciência da Informação pelo PPGCI/UFBA. Bolsista de Iniciação Científica em Pesquisa (PIBIC) no período 2007/2008. (igor.guimaraes@hotmail.com)

Analisando este fato dois pontos se destacam como pano de fundo: o primeiro é o empenho da Biblioteconomia, da CI e da TI em promover a acessibilidade à informação; e o segundo, cuja razão de ser está inter-relacionada com o primeiro, é permitir que o conhecimento aflore, tendo em vista que este é o construtor de todos os avanços econômicos, sociais, científicos e tecnológicos até então alcançados pelo homem.

Na verdade, o que se evidencia é a contínua procura pela acessibilidade do conhecimento, desde as bibliotecas tradicionais aos repositórios digitais, cujas funções de armazenamento, preservação e disseminação da informação são coincidentes e visam, em última instância, a permitir o acesso do sujeito à produção intelectual da humanidade.

Pode-se observar que a questão não é apenas tecnológica, mas nela estão também implícitas as conquistas sociais, obtidas pela socialização do conhecimento, resultante da política de democratização da educação, da criação das bibliotecas e arquivos públicos, dos museus, da internet e, por fim, da web 2, que ao introduzir recursos que possibilitam a colaboração dos diversos atores sociais via rede, tais como as tecnologias wiki, o twitter, o MSN, os fóruns de discussão, etc. trouxe novas formas de aprendizagem e permitiram a construção do conhecimento de modo colaborativo, quebrando a unilateralidade e a hegemonia científica na construção do conhecimento.

Neste percurso de universalização do conhecimento houve um movimento de autonomia do usuário na busca da informação, que exige a aplicação de trajetórias cognitivas mais complexas para tornar efetivo o processo de busca e uso da informação. A resposta dada por parte das Tecnologias da Informação e Comunicação (TICs) foi a construção de soluções interativas e amigáveis e, pelas bibliotecas, um movimento de capacitação do usuário – a *information litteracy,* o que envolve mediação e aplicação de teorias cognitivas e a implantação de serviços orientados pela filosofia da interatividade entre as partes e a colaboração simultânea, que têm sido identificados como Biblioteca 2.0.

Procura-se refletir, neste artigo, sobre o movimento deliberado das tecnologias colaborativas e de acesso livre, no sentido de criar facilidades de acesso e uso da informação para o usuário, fazendo

surgir meios mais flexíveis, atrativos e dinâmicos. No bojo desta proposta das TICs está implícita a necessidade do aprofundamento de estudos sobre o usuário e sua complexa manobra cognitiva; sobre o emaranhado de fontes de informação advindas do avanço tecnológico; sobre o desempenho da biblioteca na aplicação destas inovações científicas e tecnológicas; e sobre o papel mediador do bibliotecário neste contexto, de sofisticados mecanismos e ferramentas de busca em razão das múltiplas sintaxes.

EM BUSCA DA ACESSIBILIDADE: UM POUCO DE HISTÓRIA

As bibliotecas e arquivos foram os primeiros sistemas de informação estruturados por homens de saber na perspectiva de permitir o acesso à informação. Desde a Antiguidade até a Idade Contemporânea o papel das bibliotecas tem sido universalizar o conhecimento produzido e acumulado pela humanidade, transmitindo-o a todas as gerações. Para isto grandes bibliotecas foram criadas, como a Biblioteca Grega de Alexandria, no Egito, a Biblioteca de Nínive, na Babilônia, A Biblioteca Nacional da França, a Biblioteca do Congresso, em Washington, a Biblioteca Nacional, no Rio de Janeiro, e inúmeras outras. Todas, porém, motivadas por um único princípio: a acessibilidade da informação.

Cedo, no entanto, os primeiros bibliotecários descobriram que não basta reunir o acervo, é preciso organizá-lo, para permitir o acesso à informação nele registrada. Nesta perspectiva, Calímaco (c305–c.240 a. C.), poeta e bibliotecário grego, produziu em 120 rolos de papiro o Pinakes, catálogo que referencia grande parte do acervo da biblioteca de Alexandria; e Conrad Gesner (1516–1565), bibliotecário suíço, organizou e publicou a bibliografia *Bibliotheca Universalis,* isto para citar alguns exemplos.

Sobre este assunto, Prado recorre ao discurso poético de Quintana (1993) para explicar o papel da biblioteca na sociedade:

> A biblioteca não resolve qualquer problema da nossa cultura,

> não muda o mundo, mas guarda, cuida, organiza e classifica livros que também não mudam a estrutura social, mas fornecem informações construídas socialmente que podem mudar as visões de mundo das pessoas que os lêem. (PRADO, 2009)

A invenção da imprensa móvel, por Gutenberg e a contínua melhoria dos processos de registro do conhecimento ampliaram muito a produção de livros, enquanto que as revoluções científicas e industriais impulsionaram o desenvolvimento do conhecimento e fizeram surgir os periódicos científicos especializados. Livros, periódicos, jornais etc. passaram a ser fonte de conhecimento e a Biblioteconomia foi levada a desenvolver métodos e técnicas mais eficientes para coletar, organizar e disseminar o acervo produzido pela Humanidade.

Para permitir o acesso ao conhecimento foi preciso catalogar, classificar e indexar as publicações, de modo a representar o conteúdo nelas registrado. Pensadores como Bliss, Melvil Dewey, Ranganathan, Paul Otlet, Henri La Fontaine e outros criaram as classificações de assunto, baseadas nas classificações filosóficas do conhecimento. As grandes bibliotecas criaram suas próprias regras de catalogação, que passaram a ser adotadas também por outras, como foi o caso do Código de Catalogação da Biblioteca Vaticana e, mais tarde, o Código de Catalogação Anglo-americano. Para a indexação, surgiram listas de cabeçalhos de assunto como a Lista de Sears, a da Biblioteca do Congresso americano, a do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict), e os vocabulários controlados, ou seja, Thesauri gerais, como o da Unesco, e especializados em várias áreas do conhecimento, como o da Organização para a Cooperação e Desenvolvimento Econômico (OCDE).

Todo este arsenal de métodos e técnicas representou momentos diferentes de evolução da Biblioteconomia, na tentativa de bem-representar o conhecimento, a fim de facilitar a recuperação da informação e sua disseminação, ou seja, promover o acesso amplo e irrestrito ao conhecimento. O fenômeno da explosão da informação, motivado pelo incremento da produção científica nas duas guerras mundiais, levou à institucionalização do Controle Bibliográfico Universal, programa que tem o fim de manter o registro da produção editorial dos países

e tornar esta produção acessível mundialmente, por meio de um código universal, o International Standart Book Number (ISBN), para livros, e o International Standart Serial Number (ISSN), para periódicos. Também a criação de Sistemas de Recuperação da Informação, baseados na indexação de periódicos, ampliaram a possibilidade de acesso ao documento.

O advento da informática veio otimizar os processos da Biblioteconomia, com a transformação dos catálogos de consulta em bancos de dados bibliográficos e, por fim, o surgimento das redes digitais, Internet e intranets, possibilitaram a ampla divulgação dos acervos das bibliotecas para o mundo e no espaço das organizações. Assim, os bancos de dados puderam ser disponibilizados nas redes e acessados a distância e as bibliotecas tornaram-se digitais ou virtuais. Mas o princípio de ordenamento da informação continua baseando-se na dinâmica do seu fluxo. O professor Luís Milanesi (2002), um dos luminares da Biblioteconomia brasileira, diz "O Homem registra para reter, e o registrado não-encontrável, na prática, é igual ao inexistente". Por isto, os pilares da Biblioteconomia se assentam nos procedimentos de coleta, processamento, recuperação e disseminação da informação. E todo seu movimento, no tempo, tem sido em busca de novos procedimentos e recursos para bem-representar a informação e torná-la acessível e útil ao usuário, no momento oportuno.

Mey (2009) assevera que os princípios da Biblioteconomia têm permanecido basicamente os mesmos, durante milênios, porém o arsenal técnico, equipamentos e mecanismos que esta ciência utiliza, para tornar disponível aos usuários a informação, tem sofrido profundas modificações e, em sua reflexão, a autora delineia a estrutura básica da disciplina Biblioteconomia:

> [...] a) aquisição dos registros do conhecimento (desde a coleta de materiais tangíveis à criação de acervos digitais); b) organização dos registros do conhecimento (inclui análise, representação, criação de instrumentos de análise e representação, estudos teóricos e práticos, os mais variados); c) disseminação dos registros do conhecimento (abarca tanto os instrumentos de dis-

> seminação como os estudos sobre os usuários dos registros do conhecimento e outros aspectos teóricos e práticos). (MEY, 2009)

Avaliando esta assertiva e o percurso que têm seguido as Tecnologias de Comunicação e Informação (TICs) pode-se afirmar que elas representaram um avanço incomensurável no que tange aos processos de registro, tratamento e disseminação da informação, tornando-se uma forte aliada da Biblioteconomia, até uma certa fase sem nenhum conflito.

No entanto, os avanços das TICs na direção de soluções tecnológicas com capacidade de armazenamento ilimitado, com muitas funcionalidades e focadas na interatividade e hipertextualidade, tais como os bancos de dados de textos completos, as bibliotecas virtuais, os recursos de hipertexto, os navegadores, os motores de busca e finalmente o surgimento da Web 2.0, com as tecnologias colaborativas e a filosofia do acesso aberto, inundaram a Internet com as mais variadas fontes de informação e caminharam cada vez mais na direção da autonomia do usuário. Nesta linha de pensamento, Maruyama (1965) afirma:

> A tecnologia de armazenamento da informação própria da internet nos aponta para novas maneiras de armazenar; organizar e buscar a informação desejada. Os hipertextos organizam a informação onde os textos se repetem mutuamente. Um universo relacional em contraposição a um universo classificador já havia sido vislumbrado por filósofos e cientistas (MARUYAMA, 1965).

Ainda neste sentido, Dahlberg (1993) observa que os sistemas de organização do conhecimento bibliográfico e os alfabéticos aplicados pelos bibliotecários para representar e organizar o conhecimento foram ignorados pelos especialistas das tecnologias digitais, sob a justificativa de que sendo o conhecimento tão complexo, não poderia ser abarcado por teorias e técnicas tão genéricas.

Frente ao vasto conjunto de funções que o contexto sociocognitivo e tecnológico apresenta, deve-se refletir se a biblioteca e seus profissionais se prepararam para atuar na complexa rede social e colaborativa

criada com a Internet, aplicando as tecnologias de ponta voltadas ao armazenamento, representação, recuperação e disseminação da informação, dentro da lógica de construção, desconstrução e reconstrução do conhecimento advinda da Web.

RUMOS DOS PROCESSOS E TÉCNICAS DE REGISTRO E DISSEMINAÇÃO DO CONHECIMENTO

A legitimação da ciência se dá por meio da comunicação científica, que além de permitir a circulação do conhecimento e sua interação entre os pares assegura também seu registro e recuperação histórica.

Os investimentos feitos em TI têm resultado na diversificação contínua de meios e suportes voltados à disseminação da produção científica, promovendo os princípios básicos da comunicação científica: registro, disseminação, fidedignidade e acessibilidade. O advento da internet, além de ampliar as possibilidades de acesso às fontes de informação, introduziu novas estratégias de comunicação entre a comunidade científica.

A Iniciativa dos Arquivos Abertos – Open Archives Initiative (OAI) –, por exemplo, que surgem como uma tentativa na solução para o alto custo das assinaturas de periódicos científicos, a lentidão na disseminação da produção científica, as barreiras impostas pela legislação de direitos autorais; e a necessidade de reduzir o fosso científico, tecnológico e cultural que se instalou entre os países ricos, emergentes e pobres, em função das disparidades educacionais e culturais, para as quais contribuem a centralidade, nas regiões mais prósperas, da produção e distribuição do conhecimento, reduzindo a possibilidade do sujeito, das regiões periféricas, ampliar suas competências cognitivas e de aprendizagem, situação que contribui significativamente para acirrar o quadro de exclusão e desigualdade social.

Ferreira considera que a iniciativa de arquivos abertos representou:

> Soluções técnicas efetivas, ágeis, econômicas e viáveis para que as comunidades científicas reconstruam práticas e processos

> de comunicação científica, sistemas de gestão cooperativos, mecanismos de controle bibliográfico, preservação da memória, promovendo a consolidação de seus *corpus* de conhecimento. (FERREIRA, 2007)

A inserção de tecnologias colaborativas a partir da criação da Web 2.0 promoveu mudanças nas relações de produção/geração, disseminação e acesso à informação, a exemplo do Modelo de Comunicação Todos-todos (MENDONÇA, 2007), baseado na filosofia participativa "de muitos para muitos", em que todos os gêneros de comunicação são híbridos e dinâmicos na construção de registros, enquanto os documentos tornam-se, muitas vezes, produtos coletivos, reconhecidos socialmente.

Este rearranjo no fluxo da informação tornou mais efetivo o processo de comunicação entre os pares e entre estes e a sociedade sendo que, com a intermediação deste aparato tecnológico, a disseminação da informação agora pode-se fazer em tempo real, de modo compartilhado e interativo. Isto porque, a partir destes avanços, à comunicação científica é incorporado o uso de recursos eletrônicos como e-mail, listas de discussão, *chats,* catálogos *on-line,* boletins, *blogs, twitter,* entre outros, facilitando a troca de ideias e aconselhamentos entre os pares da comunidade científica, prática que existe entre os pesquisadores desde que a ciência ganhou corpo, mas que se inicia, de forma lenta, com as cartas, que, aliás, foram a origem dos periódicos científicos.

Neste afã de democratização do conhecimento, Paul Ginsparg, do Laboratório Nacional de Los Alamos, Novo México, teve a iniciativa de criar o *arXiv,* um repositório digital de livre acesso, considerado o primeiro repositório *e-print,* que permitiu aos pesquisadores o auto-arquivamento a distancia, de sua produção científica. (BAPTISTA, 2007) Desde então, o Movimento de Acesso Livre mobilizou a comunidade científica no sentido de criar uma mentalidade de abertura ampla e irrestrita na divulgação do conhecimento científico, quebrando as barreiras que o protegiam e incrementando soluções técnicas que terminaram por viabilizar a tecnologia dos Open Archives.

Este trabalho cooperativo possibilitou a ampliação do diálogo entre os atores sociais, entrelaçando o público e o privado, a comunidade científica e leiga, governos e cidadãos; a relação política entre o acesso à informação e os direitos humanos; esta abertura passou a exigir ações mais rigorosas no controle bibliográfico, por meio dos recursos de metadados e interoperabilidade e preservação digital. (WEITZEL, 2006)

A criação de soluções que facilitem o acesso do usuário ao conhecimento guarda uma relação direta com a autonomia deste no manuseio dos recursos informacionais. Neste sentido, as TICs inovaram criando ferramentas amigáveis e colaborativas, que interagem com o usuário no processo de busca, permitem que ele próprio registre sua produção científica na rede e ainda possibilitam-lhe a oportunidade de dialogar com seus pares e com os diversos atores da rede. Ou seja, o processo de informação e comunicação passou a ser exercido pelas próprias ferramentas, o que tem sido identificado como "processos de comunicação mediados por computador". (BASTOS; OLIVEIRA; VIDOTTI, 2009) A plataforma Web 2.0 é a mais recente concretização das TICs em torno da autonomia do usuário no manejo das ferramentas digitais; e este fenômeno informacional já ganhou o título de "desmediação".

Nesta linha estão os repositórios digitais, ferramentas incrementadas pela política de acesso livre à produção científica e pela tecnologia de open archives, que foram também impulsionadas pela aceitação das instituições científicas, tendo em vista que têm como finalidade reunir, preservar, dar acesso e disseminar o conhecimento produzido pela comunidade científica de uma determinada instituição, contribuindo para o reconhecimento e visibilidade do pesquisador e de sua instituição de origem, bem como para acelerar o avanço da ciência.

Nos caminhos trilhados pelas TICs estão também as bibliotecas, consideradas desde a Antiguidade, e citadas em várias literaturas, como repositórios do conhecimento e da cultura, tendo em vista que têm a missão de reunir, arrolar, armazenar um acervo com a finalidade de organizar e disseminar a informação.

A BIBLIOTECA E OS BIBLIOTECÁRIOS EM TEMPOS DE WEB 2.0: CONFIGURANDO A BIBLIOTECA 2.0

A lógica da acessibilidade e da abrangência que sustenta a rede fez dela, em um determinado momento, uma aliada da biblioteca, porém o grau de autonomia que ganhou com o enriquecimento dos conteúdos em sua teia e a diversidade de ambientes informacionais com funções múltiplas, como, por exemplo, os periódicos *on-line,* a digitalização de textos retrospectivos na íntegra, o livro eletrônico, os portais corporativos e inúmeras outras fontes de informação, agora nela disponíveis, de modo portátil, com possibilidade de recuperação ágil e fácil, levaram-na a se tornar uma concorrente das bibliotecas, sem deixar de ser, ao mesmo tempo uma ferramenta auxiliar indispensável ao efetivo desempenho das bibliotecas contemporâneas.

A rede é uma concorrente com muitos atrativos, um deles, a onipresença, determinante na escolha do usuário, considerando que os estudos sempre apontaram que a preferência do usuário na busca da informação recai na facilidade de acesso. A rede representa a teia que Ptolomeu tentou construir com a biblioteca de Alexandria e Paul Otlet, com o Mundaneum, ou seja, ela é a concretização do desejo do homem, na perspectiva de reunir e compartilhar, agora colaborativamente, o conhecimento produzido por todos, tendo em vista que as ferramentas tecnológicas permitem esta façanha. Bastos, Oliveira e Vidotti (2009) apontam como resultado de pesquisa realizada sobre a Web 2.0:

> A inserção dessas tecnologias apresenta-se como inovação, mas sua natureza está vinculada à tradição e à missão das bibliotecas e dos repositórios. Os ambientes científicos informacionais observados que apresentaram as tecnologias da web 2.0 estão estruturados de uma forma mais flexível, atrativa e dinâmica, no qual o usuário se torna o ator principal na construção de seu ambiente, possibilitado pelos recursos de customização e personalização. (BASTOS; OLIVEIRA; VIDOTTI, 2009)

Por outro lado, os processos técnicos biblioteconômicos voltados

à recuperação da informação sempre privilegiaram os princípios da classificação do conhecimento e indexação dos conteúdos, daí a permanente busca de aperfeiçoamento nestes processos até o surgimento de ferramentas reconhecidas e utilizadas universalmente, como a Classificação Decimal de Dewey (CDD) e a Classificação Decimal Universal (CDU) e os thesauri. A rede, no entanto, conforme já foi dito neste texto por meio das palavras de Maruyama (1965), estrutura-se com uma teia relacional, que permite o entrelaçamento de conteúdos, maximizando as possibilidades de busca e recuperação, haja vista resultados de buscas feitas na Internet com a mediação de motores de busca, como Yahoo, Google etc., em que se encontram preciosidades, mas também respostas sem nenhum significado para o usuário.

Há que se considerar que, em um ambiente informacional científico, uma das prioridades das ferramentas de busca deve ser garantir a relevância dos resultados da pesquisa bibliográfica, a fim de assegurar ao pesquisador a entrega dos mais significativos documentos de sua área de interesse. Estudiosos de renome do campo da indexação e recuperação da informação, tais como, Lancaster (2004) e Saracevic (2009) têm ressaltado este ponto, ou seja, da necessidade de que princípios como relevância e precisão, sempre utilizados em sistemas de recuperação da Informação científicos, sejam preservados como forma de garantir a qualidade nas pesquisas bibliográficas.

A literatura, contudo, vem mostrando que há um forte movimento, envolvendo bibliotecas e bibliotecários, voltado à aplicação da Web 2.0 como plataforma para divulgação de seus serviços, como forma de levar a biblioteca até o usuário. Este movimento é identificado como Biblioteca 2.0 e o princípio que o sustenta é a possibilidade de incrementar os serviços da biblioteca, introduzindo ferramentas tecnológicas colaborativas, que lhe permitam a criação de serviços e produtos informacionais interativos, concretizando a realização de ações e serviços com o compartilhamento do usuário.

> O coração da biblioteca 2.0 é a mudança centrada no usuário. É um modelo que visa a implementar serviços de biblioteca, na perspectiva da mudança intencional e constante, com a partici-

> pação do usuário na criação e transformação dos serviços, tanto virtuais quanto físicos, levando sempre em conta o processo de avaliação. O objetivo é atender bem os clientes atuais, com a melhoria e oferta orientada de serviços, bem como alcançar novos usuários. Cada componente do processo por si só é um passo na direção de atender às demandas dos usuários, contudo, é através da aplicação combinada de todos estes processos e recursos que se pode alcançar o propósito da Biblioteca 2.0. (CASEY; SAVASTINUK, 2006)

Dentre estes recursos, estão os sistemas de utilização de tag, ou seja, o "tagueamento", que é a adição, pelo usuário, de palavras-chave livres a recursos da Web, tais como *websites,* páginas, imagens, músicas etc. Segundo Bastos, Oliveira e Vidotti (2009), o tagueamento livre é um sistema de folksonomia, porque os termos são introduzidos por qualquer pessoa usuária da rede, diferindo das ordenações fixas e predefinidas da taxonomia e têm a finalidade de possibilitar o diálogo entre comunidades de interesse, tendo em vista que se adaptam ao fluxo dos discursos.

É oportuno, no entanto, refletir sobre a afirmação destes autores acima citados de que este é um processo alternativo, que estimula a conversação coloquial entre os atores da rede, não um modelo credenciado e avalizado de armazenar e recuperar informações científico-tecnológicas, tais como os grandes sistemas de recuperação da informação, a exemplo do Index Medicus, Chemical Abstracts Services e inúmeros outros com cunho científico.

> Como se observa, a escrita coletiva ciberespacial e o processo de 'tagueamento' demonstram que a abertura para o trabalho colaborativo oferece uma dinâmica alternativa (não uma substituição) ao modelo de produção, indexação, armazenamento e controle por equipe de autoridades. A partir de recursos da web 2.0, potencializa-se a livre criação e a organização distribuída de informações compartilhadas através de associações mentais. (BASTOS; OLIVEIRA; VIDOTTI, 2009)

Isto, contudo, não anula a possibilidade das bibliotecas usufruírem dos dinâmicos recursos da rede, na perspectiva da disseminação de informação e da agregação de meios que incrementam a construção do conhecimento. Spyers (2007 apud ALCARÁ; CURTY, 2009) destaca o caráter popular dos blogs como diários pessoais, mas conclui que, a facilidade com que são manejadas e a rapidez de publicação destas ferramentas favoreceram o surgimento de blogs dedicados a conteúdos científicos.

Alcará e Curty (2009) confirmam esta declaração dizendo:

> [...], – muitos profissionais e pesquisadores de importante representação na área em que participam estão se utilizando deles como uma alternativa prática e ágil para acelerar as discussões e os questionamentos frequentemente suscitados pelos seus pares, compondo, dessa forma, uma verdadeira vitrine virtual de informações científicas.

No entanto, as citadas autoras, ponderando sobre as características dos blogs concluem que, em virtude de sua natureza como mecanismo informal de comunicação, ainda que produzam e divulguem conteúdos científicos, é precipitado considerar os blogs como fontes de informação científica, antes de uma discussão sobre a qualidade das informações divulgadas neles, no entanto, dizem as autoras "Em que pese a isso, certamente o status do *blog* como um possível canal de comunicação informal em ambiente web no processo de discussão da ciência já não pode ser questionado" (ALCARÁ; CURTY, 2009), até porque estes blogs, por sua origem, quase sempre criados e administrados por pesquisadores, obtêm, naturalmente, confiabilidade e legitimidade da comunidade científica. A prova disto é que, à semelhança do ISSN e ISBN, já foi criado o International Blog Serial Number (IBSN), um meio de identificação dos *blogs* e também o Anel de Blogs Científicos (ABC), do Laboratório de Divulgação Científica (LDC) da USP.

Em 2005, Miller vê na Web 2.0 a possibilidade de a instituição biblioteca dar um salto de qualidade, tanto no que se refere à organização da informação, na medida em que pode lançar mão da Web 2.0

como plataforma e usar seus padrões, como pela possibilidade de usar seus recursos de interoperabilidade entre sistemas, o que lhe possibilita criar infinitos elos, aumentando a exposição e popularização de seus produtos e barateando seu custo, o que lhe insere no fenômeno estatístico identificado como "cauda longa" (long tail), que, além da ressonância social, é economicamente rentável.

Miller lembra que os usuários finais, tais como os blogueiros, têm sabido se aproveitar destas vantagens, mas

> [...] existe, igualmente, a possibilidade de participação das bibliotecas e organizações associadas. Bibliotecas, fornecedores de sistemas, editores, organismos normalizadores, agências governamentais precisam trabalhar juntos. Nenhum de nós pode fazer isto sozinho e todos nós podemos nos beneficiar se fizermos juntos. Podemos dar o salto? Alguém corajoso o suficiente para tentar, por favor, entre em contato. (MILLER, 2005)

Estudos posteriores demonstram que a biblioteca vem procurando transcender ao seu espaço e seu modo de funcionamento, inserindo-se nas redes sociais, reafirmando sua natureza como meio de comunicação de massa, de finalidade educativo-cultural; esta ressignificação de sua missão vem pela via da otimização de seus serviços, com a aplicação dos princípios e das ferramentas da Web 2.0, a fim de alcançar, não só os usuários cativos, mas também aqueles potenciais, que não se manifestam face a face, mas usam os recursos informativos na "sombra da rede". Rodriguez Palchevich (2009) e Murley (2008) discriminam várias ferramentas que já estão ou podem ser usadas pelas bibliotecas para estenderem o alcance de seus serviços, tais como *blogs, feeds, wikis, Instant Messaging* (IM), *chats, device-independent, Websites,* fóruns de discussão etc., a fim de "fomentar a colaboração e o intercâmbio ágil de informação entre os usuários". (RODRIGUEZ PALCHEVICH, 2009)

A bibliotecária alemã Heike Neuroth (2009), embora tivesse questionado se, neste contexto da Web 2.0, as bibliotecas ainda desempenham um papel central no que se refere ao gerenciamento e

difusão da informação, ela própria acena com as possibilidades que as bibliotecas têm de ampliar seus serviços com estas ferramentas e exemplifica com a biblioteca alemã vascoda, que é um portal de informações científicas abrangendo todas as áreas do conhecimento, criado a partir da fusão de várias instituições e redes de informação e ainda algumas bibliotecas virtuais especializadas, colocando à disposição do usuário a possibilidade de busca interdisciplinar e de navegação em numerosos portais especializados. E a europeana, constituída a partir das coleções de bibliotecas, museus e arquivos da Europa, dentre eles o Rijksmuseum, em Amsterdã, a British Library, em Londres e o Louvre, em Paris, a qual já conta com mais de quatro milhões de itens digitais, reunindo imagens, pinturas, desenhos, mapas, fotografias, livros, textos, jornais, cartas, diários, registros sonoros, vídeos, filmes e emissões de televisão.

No Brasil, um exemplo de incorporação intensa de tecnologia que pode ser citado é o da Biblioteca Nacional de Brasília, cujo projeto privilegiou o modelo de biblioteca híbrida, ou seja, um espaço constituído de acervo físico, mas igualmente de recursos e serviços digitais tais como, computadores, Web pública, espaços apropriados para transmissão de vídeos e videoconferência etc., tudo isto posto à disposição do público.

Refletindo sobre este assunto, Miranda, Leite e Suaiden (2007) denotam que as transformações sociais e o advento da tecnologia promoveram a transformação dos próprios princípios da biblioteca. Eles destacam o papel histórico e, porque não dizer heróico, das bibliotecas nacionais de diversos países, no sentido de preservar o patrimônio bibliográfico nacional, de criar o programa de controle bibliográfico nacional, em consonância com o programa de controle universal, e de incentivar a implantação de sistemas nacionais de bibliotecas públicas e confrontam estes objetivos com o objetivo introduzido pelas bibliotecas híbridas, qual seja o de intensificar a acessibilidade documentária em rede, para isto estas bibliotecas ancoram seus serviços na aplicação das TICs, quer na realização dos processos-meio, quer na concretização de suas atividades finalísticas, agora mais facilitadas pelas ferramentas da Web 2.0. Eis um trecho de artigo científico dos

citados autores a respeito desta temática:

> O desenvolvimento das (então denominadas) "novas tecnologias da informação" permitiu a montagem de complexos bancos de dados, a diversidade fantástica de novos suportes informacionais (da microficha aos repositórios digitais dos nossos dias), além do surgimento da *web* e da Internet, que mudariam completamente os paradigmas da formação e desenvolvimento de acervos. Também transformaram as instituições devotadas à seleção, aquisição, organização, difusão e preservação de uma agora inimaginável quantidade de conteúdos de informação de todo o tipo em línguas, níveis de leitura e em formatos convencionais e multimídia, de forma hipertextual e instantânea. (MIRANDA; LEITE; SUAIDEN, 2007)

Segundo os autores a disponibilidade e a acessibilidade documentária são dois princípios básicos que norteiam a missão e o *modus operandi* das bibliotecas, sendo que a oferta de serviços de informação via Web está fortemente relacionada com a promoção da acessibilidade, porém, conforme eles asseguram, este fenômeno só acontece a partir da disponibilidade. O professor Miranda (2005 apud MIRANDA; LEITE; SUAIDEN, 2007), assim, conceitua os dois princípios:

> A capacidade que as bibliotecas têm de selecionar, adquirir, organizar e prestar serviços a partir de uma coleção física de documentos é tradicionalmente denominada **disponibilidade documentária.** A capacidade que as bibliotecas têm de organizar serviços de busca de documentos e informações em outros repertórios para o atendimento de seus usuários está baseada na **acessibilidade documentária.** (MIRANDA, 2005 apud MIRANDA; LEITE; SUAIDEN, 2007)

De fato, diante deste emaranhado de recursos e ferramentas informacionais e das constantes inovações aportadas pelas TICs a situação futura das bibliotecas, assim como de todos os segmentos mais

intimamente relacionados com o processamento de informação, é nebuloso, mas também promissor, conforme é visto em todos os exemplos citados. As instituições, os profissionais da área, a própria sociedade, no entanto, cobram o alinhamento dos serviços prestados pela biblioteca às exigências da vida moderna, que se fundamenta, sobretudo, na ampla oferta ao usuário das facilidades de acesso à informação. Neste sentido, conforme Neuroth (2009), a biblioteca deve incorporar a ideia de "um acesso homogêneo a muitas ofertas heterogêneas", ou seja os serviços bibliotecários devem centralizar variadas funções de busca e atendimento ao usuário, contemplando questões científicas e interdisciplinares.

Andrade (2009) explica que a tendência é que os diversos recursos se integrem se complementando e que os *blogs* e os agregadores de notícias (RSS) são muito mais usados pelas bibliotecas que os recursos para criação compartilhada, os *wikis;* Arnal (2008 apud ANDRADE, 2009) e Mannes (2007 apud ANDRADE, 2009) esclarecem que na Europa as bibliotecas já se conscientizaram da importância da aplicação destes recursos e já os usam largamente; Andrade (2009), no entanto, constata em sua pesquisa o pouco uso destas tecnologias pelas bibliotecas físicas brasileiras, enquanto que, nas bibliotecas virtuais, este uso já é um fato consumado.

Na perspectiva de estimular e orientar a biblioteca a enfrentar os desafios da contemporaneidade, Neuroth (2009) apresenta as 12 teses promulgadas pela Comissão Consultiva de Biblioteconomia da ekz para a biblioteca do futuro, quais sejam:

> 1. As bibliotecas devem se colocar de forma proativa frente ao desafio de ser "um prestador de serviços digitais".
> 2. As bibliotecas repousam sobre dois pilares, um com paredes e teto, e outro virtual.
> 3. As bibliotecas devem ser curiosas, assumir um papel pioneiro na rede e utilizar as plataformas disponíveis.
> 4. As bibliotecas devem se tornar decisivas, já que a tomada de decisões e a definição de rumos no mundodigital precisam ocorrer de forma cada vez mais rápida e num prazo mais

curto do que nunca. Só desta formaelas poderão determinar os acontecimentos na rápida mutação das mídias, através de ofertas atualizadas.
5. O seu papel como „bem público“ deve ser incutido e expandido no inconsciente coletivo.
6. Os políticos devem ser convencidos da necessidade de que as bibliotecas ofereçam produtos digitais.
7. As ofertas digitais devem ser específicas para cada Público-alvo, e também organizadas de forma intercultural.
As bibliotecas devem contribuir para criar um justo equilíbrio entre fornecedores de conteúdos, mediadores e clientes.
9. Um grande tesouro das bibliotecas é o seu conhecimento sobre os interesses dos clientes, que pode ser utilizado de forma intensiva e ao mesmo tempo responsável.
10. As bibliotecas precisam colocar-se, de forma mais intensa, como fornecedora de conhecimento e local de aprendizado. (E-Learning).
11. As bibliotecas devem se apresentar na rede como parceiras dos movimentos sociais (wikipedia, social tagging, etc).
12. Para os bibliotecários, isto significa continuar a desenvolver, de forma sistemática, as suas competências no mundo digital.
http://bibliothekarisch.de/blog/2008/07/31/ekz-bibliothekarischerbeirat-stellt-12-thesen-zur-bibliothek-der

Apesar da complexidade das funções que têm assumido, as bibliotecas guardam sua missão básica de reunir, preservar, ordenar para tornar disponível o patrimônio científico-cultural de uma organização, tal como, em tempos remotos, a biblioteca especializada e universitária, a exemplo da biblioteca nacional que reúne o acervo de um país, arrolavam a chamada “Memória Científica e Técnica” de uma instituição; deste modo, é lícito considerar que as bibliotecas devem assumir a liderança de planejamento, implementação e gerenciamento dos repositórios institucionais, em vista da natureza destas fontes informacionais. Pesquisa realizada por Maria Carmen Romcy

de Carvalho (2009) relativa à iniciativa dos repositórios em bibliotecas universitárias aponta que, em sua maioria, no Brasil, os repositórios de universidades são gestados por suas bibliotecas; no entanto, o fato de não haver unanimidade, demonstra que não há um consenso sobre isto. Cabe, no entanto, à biblioteca a iniciativa de legitimar este direito com ação empreendedora, competência e criatividade.

Quanto ao fenômeno da desmediação, pode-se dizer que ele se coaduna com os objetivos da biblioteca de dar autonomia ao usuário na busca e consulta das fontes de informação e que estão relacionados com as ações de "educação ou instrução do usuário", realizadas pelas bibliotecas desde o século passado e que mais recentemente ressurgiu com o nome de *information litteracy,* ou letramento informacional, ou ainda desenvolvimento de competência informacional, tema a ser abordado ainda neste capítulo.

Quanto à capacitação do bibliotecário para atuar neste contexto de uso intensivo de tecnologia da informação, pode-se afirmar que muitos deles manejam com desenvoltura as soluções tecnológicas e eles podem ser encontrados e contatados no pleno movimento dialógico da rede, no entanto, pressupõe-se que esta não é a tônica, além do que, é preciso que treinamentos sistemáticos sejam oferecidos para profissionais já atuantes no mercado, bem como intensificar a oferta de disciplinas de TI nos cursos regulares de Biblioteconomia, de modo a se potencializar a formação de bibliotecários com bom desempenho na geração, uso e manuseio de fontes de informação mediadas pelas TICs.

No que concerne ao relacionamento do bibliotecário com o usuário e deste com o manejo das fontes de informação estão implícitos processos intelectuais e cognitivos, evidenciando-se, assim, que a formação deste profissional deve contemplar também matérias relativas à cognição, até porque, espera-se que ele, atuando, como mediador do conhecimento, numa esfera educativa e cultural, esteja apto a construir significado e a transcender, na ação de transferência da informação, num processo dialógico e de aprendizagem recíproca.

A COGNIÇÃO COMO SUPORTE NA CONSTRUÇÃO DOS PROCESSOS DE RECUPERAÇÃO E USO DA INFORMAÇÃO

Os estudos teóricos e experimentais conduzidos a partir dos anos 50 do Século XX na perspectiva de solucionar as questões de registro e recuperação da informação, segundo Saracevic, levaram a esclarecer:

> a natureza da informação, a estrutura do conhecimento e seus registros(incluindo bibliometria), o uso e os usuários, levando a estudos do comportamento humano frente à informação; a interação homem-computador, com ênfase no lado humano da equação; relevância, utilidade, obsolescência e outros atributos do uso da informação juntamente com medidas e métodos de avaliação dos sistemas de recuperação da informação; economia, impacto e valor da informação, dentre outros. (SARACEVIC, 1996)

No bojo destes estudos Mooers (1951 apud SARACEVIC, 1996), cunhou a expressão "Recuperação da Informação" e buscava respostas para as seguintes questões:

a) como descrever intelectualmente a informação?

b) como especificar intelectualmente a busca?

c) que sistemas, técnicas ou máquinas devem ser empregados?

Estas questões são basilares na descoberta de soluções teórico-práticas relativas à recuperação da informação, tendo por isto gerado uma grande variedade de conceitos e construtos teóricos, empíricos e pragmáticos, bem como numerosas realizações práticas, daí vieram as marcantes evoluções de sistemas, técnicas e/ou máquinas utilizadas pelos processos de recuperação da informação que vão dos meios mais rudimentares às interfaces inteligentes e aos processos altamente interativos dos dias atuais.

Neste contexto, antes de tudo, quem é que vivencia os desafios da busca e do uso do conhecimento?

O EU E O NÓS NA BUSCA DO CONHECIMENTO

O fenômeno da explosão informacional fez com que os seres humanos que vivem o século XXI passassem a abordar a realidade por meio de ferramentas cognitivas, a construir representações mentais, a desenvolver competências e habilidades e refletir em torno da apreensão e da compreensão da informação. A falta de domínio das habilidades de leitura e de acesso à informação, acrescida da carência de infra-estrutura tecnológica em algumas partes do mundo, distancia os historicamente excluídos dos privilegiados. Sabe-se que toda ação que recai sobre indivíduos requer uma concepção filosófica consciente e criticamente construída.

A compreensão humana chega quando se sentem e concebem os humanos como sujeitos. De acordo com Morin (1998), a noção de indivíduo/sujeito implica, simultaneamente, autonomia e dependência, sem, contudo, reduzir-se a isso. Faz-se necessário compreender o que é, fundamentalmente, a organização viva. A biologia molecular e a genética nos deram todos os elementos para compreender essa organização.

Precisa-se, antes de tudo, partir do princípio de identidade, que possibilita um tratamento objetivo a todas as operações das moléculas, das células, de um organismo policelular. No entanto, essa objetividade também tem uma dimensão subjetiva, que possibilita a autorreferência. Posso tratar a mim mesmo, referir-me a mim mesmo, porque necessito de um mínimo de objetivação de mim mesmo, uma vez que permaneço "eu", ou seja, como sujeito. (MORIN, 1998)

Com relação aos conceitos apresentados anteriormente, é necessário que seja feita uma distinção fundamental entre o si e o não-si, distinção não apenas cognitiva; mas, ao mesmo tempo, distributiva de valor: atribui-se valor ao si, mas não ocorre o mesmo em relação ao não-si. Morin lembra que "esse processo de auto-exo-referência é o que é constitutivo da identidade subjetiva. É assim que se opera a distinção entre si/não-si, mim/não-mim, entre eu e os outros eus". (MORIN, 1998)

Dessa forma, dois princípios subjetivos se associam: o princípio da exclusão e o da inclusão. Para os linguistas, qualquer um pode dizer "eu", mas ninguém pode dizê-lo por mim. Apesar de o "eu" ser a coisa mais corrente, não é, ao mesmo tempo, coisa absolutamente única. O princípio da exclusão é inseparável do princípio de inclusão, e isso torna possível integrar a objetividade a sujeitos diferentes de nós. Pode-se integrar a subjetividade pessoal a uma subjetividade mais coletiva: "nós". (MORIN, 1998)

O indivíduo/sujeito pode tomar consciência de si mesmo por meio do instrumento de objetivação que é a linguagem. Surge, então, a consciência de ser consciente e a consciência de si mesmo, de forma claramente inseparável da auto-referência e da reflexão. É na consciência, diz Morin (1998), que se objetiva a si mesmo para ressubjetivar-se num círculo recursivo incessante.

No entanto, a condição humana está marcada por duas grandes incertezas: a incerteza cognitiva e a incerteza histórica. Há três princípios de incerteza no conhecimento: **o cerebral**: o conhecimento nunca é um reflexo do real, mas sempre tradução e construção; **o físico**: o conhecimento dos fatos é sempre dependente de interpretação; **o epistemológico:** decorre da crise dos fundamentos da certeza, em filosofia – a partir de Nietzche –, depois em ciência – a partir de Bachelard e Popper. (MORIN, 2008)

DINÂMICA NA BUSCA E NO USO DA INFORMAÇÃO

Toda esta preocupação pelos estudos da condição humana e da sociedade emerge para explicar o processo de construção do conhecimento. Todo conhecimento se constitui ao mesmo tempo uma tradução e uma reconstrução, a partir de sinais, signos, símbolos, sob a forma de representações, ideias, teorias e discursos; a organização do conhecimento é realizada em função de princípios e regras, além de comportar operações de ligação (conjunção, inclusão, implicação) e de separação (diferenciação, oposição, seleção, exclusão). O processo é circular, passando da separação à ligação, da ligação à separação, e, além disso, da análise à síntese, da síntese à análise. (STERNBERG, 2000)

O processo de recuperação da informação é, também, em essência, a resposta a um ato cognitivo, que se expressa na ação de buscar a informação para atender a necessidade do usuário. Daí justifica-se o envolvimento, neste processo da Psicologia Cognitiva, que estuda o modo como as pessoas percebem, aprendem recordam e pensam a informação. (STERNBERG, 2000)

A representação do conhecimento compreende várias maneiras nas quais nossas mentes criam e modificam as estruturas mentais que representam o que conhecemos sobre o mundo externo, envolvendo tanto a forma declarativa (saber o que), como a forma não-declarativa (saber como) de conhecimento. Pela imaginação, criamos estruturas mentais que representam coisas que no momento não estão sendo percebidas pelos órgãos sensoriais. A unidade fundamental do conhecimento simbólico é o conceito, e os conceitos podem ser organizados em esquemas, que podem incluir outros esquemas, variar em aplicação e em abstracionismo, e incluir informações sobre as relações entre conceitos, atributos, contextos e conhecimento geral, bem como informações sobre relações causais. (STERNBERG, 2000)

Feuerstein (1980), em seus estudos sobre o funcionamento da mente humana, concluiu quais aspectos levariam a um processamento mental mais eficiente e que fatores poderiam interferir neste processo. Para ele, o ser pode ser percebido sob três funções básicas: **psicológica**, corresponde aos processos biológicos internos, químicos e psicológicos; **cognitivo**, diz respeito aos processos de aprendizagem, raciocínio, percepção, inteligência; **e afetiva, ou conotativa**, se refere aos aspectos emocionais e aos sentimentos. O autor constata a existência de alguns processos estruturados e dinâmicos que se combinam e se organizam de modo a fazer operar a estrutura cognitiva, o que denominou funções cognitivas. (SOUZA; DEPRESBITERIS; MACHADO, 2004)

As características fundamentais de uma função cognitiva, segundo Feuerstein (1980), são: **capacidade, necessidade e orientação**. Este conjunto dá origem a uma operação mental. A **capacidade**, caracterizada por sua dinâmica, possibilita o desempenho em seus vários níveis de complexidade e é influenciada por fatores genéticos, endógenos e externos. A **necessidade** estimula ou inibe a ação, de acordo com

seu grau de intensidade para mobilizar uma função cognitiva. A necessidade está intimamente relacionada às exigências do mundo. A **orientação** dirige a função cognitiva, determina escolhas individuais, métodos e estratégias para lidar com o problema, dirige as escolhas do indivíduo sobre o domínio do conteúdo, o ambiente ou a estrutura para a qual se dirigem os esforços para a resolução de problemas.

Conforme assinala Choo (2006), as necessidades de informação não surgem plenamente elaboradas, mas vão sendo clarificadas e definidas durante um período de tempo. Elas surgem da incerteza, a qual diminui à medida que a necessidade de informação toma a forma de questionamentos ou temas conscientes e, depois, formalizados. Os questionamentos mais constantes versam sobre o que se deseja saber, por que se precisa saber isso, qual é o problema, o que já se sabe, o que se espera descobrir, como isso vai ajudar, como se precisa saber isso, e de que forma se precisa saber isso (CHOO, 2006).

Definida a necessidade informacional, o próximo passo é o da busca de informações para atendê-la. Em seus estudos sobre o processo de busca, Kuhlthau (1993) conclui que este é um processo que se constitui de uma série de escolhas pessoais que vão determinar a eficiência do uso de fontes e estratégias. As crenças e expectativas sobre que fontes devem ser usadas e sobre a relevância ou irrelevância das fontes selecionadas baseiam-se na experiência, no potencial e no estilo cognitivo do indivíduo.

Kuhlthau (1993) divide o processo de busca da informação em seis estágios: iniciação, seleção, exploração, formulação, coleta e apresentação. Cada estágio desse processo de busca caracteriza-se pelo comportamento do usuário em três campos de experiência: o emocional (sentimentos), o cognitivo (pensamento) e o físico (ação).

Segundo Choo (2006), a busca da informação percorre a seguinte trajetória: iniciar, encadear, buscar, comparar, selecionar, extrair, verificar e finalizar. Iniciar, encadear e buscar visam a ajudar o indivíduo a desenvolver um foco para a pesquisa e uma estratégia para sua execução.

Como observou Choo (2006), a existência do vazio cognitivo impulsiona o processo de busca e é acompanhado de diferentes estados

emocionais. Os primeiros estágios da busca de informação são caracterizadas pela ansiedade, confusão, frustração e dúvida. À medida que o processo é bem sucedido, a confiança cresce e surge um sentimento de satisfação. Os estados emocionais influenciam: a) maneira como o indivíduo processa e usa a informação; b) capacidade do usuário de construir significado; c) o modo de focalizar a busca; d) capacidade de distinguir informações relevantes e irrelevantes; e) modo de lidar com o emocional e as expectativas; f) grau de interesse na pesquisa. (CHOO, 2006)

Feuerstein (1980) defende que o ato mental é produto de operações mentais e funções cognitivas que podem ser entendidas como um conjunto de ações internalizadas, organizadas e coordenadas, no que se refere às informações que recebemos de fontes internas e externas.

QUADRO 1 - FASES DO ATO MENTAL

FASE DE ENTRADA (INPUT)	FASE DE ELABORAÇÃO	FASE DE SAÍDA (OUTPUT)
• Percepção clara e precisa • Comportamento exploratório sistemática • Uso de vocabulário e conceitos apropriados • Orientação espacial eficiente • Orientação temporal eficiente • Constatação da constância e permanência do objeto • Coleta de dados com precisão e permanência do objeto • Coleta de dados com precisão e exatidão • Considerar duas ou mais fontes de informação.	• Perceber o problema e defini-lo com clareza • Facilidade para distinguir dados relevantes e irrelevantes • Exercitar a conduta comparativa • Amplitude do campo mental • Percepção global (não episódica) da realidade • Uso de raciocínio lógico • Interiorização do próprio comportamento • Exercício do pensamento hipotético-inferencial • Traçar estratégias para verificar hipóteses • Conduta planejada • Elaboração de categorias cognitivas • Aplicação da conduta somativa • Facilidade para estabelecer relações viturais.	• Comunicação descentralizada (não-egocêntrica) • Projeção de relações virtuais • Expressão sem bloqueio na comunicação de respostas • Respostas certas sem ensaio e erro • Uso de instrumentos verbais adequados • Precisão e exatidão na comunicação sem respostas • Eficácia no transporte visual • Conduta controlada, não impulsiva.

Fonte: Feurstein, R.; Feurstein, S. (1994).

As operações mentais de acordo com Feuerstein (1980) são: identificação, comparação, análise, síntese, classificação, codificação, decodificação, projeção de relações virtuais, diferenciação, representação mental, transformação mental, raciocínio divergente, raciocínio hipotético, raciocínio transitivo, raciocínio analógico, raciocínio lógico, raciocínio silogístico e raciocínio inferencial.

Numa analogia com o processamento das informações no computador, Feuerstein classificou as funções cognitivas na mente humana em três fases: entrada, elaboração e saída. Contudo, essa classificação não pode ser vista de forma estanque. As fases constituem um ato mental indissociável, no qual cada fase tem relação com as demais.

Na fase de elaboração ocorre o processamento da informação, com a busca de soluções para o problema. Na fase de saída indica-se a resposta adequada ao problema apresentado na fase de entrada, o que implica habilidades precisas de comunicação.

A busca e o uso da informação em todas as suas dimensões: acesso, análise, interpretação, avaliação, produção etc. subtende o domínio de competências e habilidades de busca da informação, conforme vem-se discutindo neste texto, mas também de leitura e compreensão, ambos relacionados com cognição.

A LEITURA E SEUS SIGNIFICADOS NO ATO DE BUSCA E DE USO DA INFORMAÇÃO

> A relação entre o pensamento e a palavra não é uma coisa, mas um processo, o movimento contínuo de vaivém do pensamento para a palavra e vice-versa [...] O pensamento não é simplesmente expresso em palavras; é por meio delas que ele passa a existir. (VYGOTSKY, 1993)

O acesso ao conhecimento tornou-se uma questão de sobrevivência, uma vez que se constitui fator de competitividade em qualquer sistema social. A queda de fronteiras entre as áreas do saber, as novas propostas metodológicas e epistemológicas, revestem de especial

importância as reflexões sobre a leitura e sobre os processos de aquisição do conhecimento.

A leitura, ou a compreensão em leitura, consiste numa atividade de processamento e integração da informação, realizada pela mente humana. O leitor chega ao texto com seu conhecimento prévio sobre o mundo e sobre a língua; Em seguida, compara a estrutura semântica às estruturas pré-existentes em sua memória e busca a integração dessas estruturas, ou seja, busca a compreensão. (GABRIEL, 2005)

Nem sempre a compreensão se dá na plenitude, por isso fala-se em níveis de construção do sentido, pois os esquemas cognitivos são particulares a cada indivíduo, construídos pelas experiências vivenciadas. As conexões entre os conhecimentos adquiridos ao longo dos anos vão sendo estruturadas continuamente, de uma forma que é particular a cada um. (GABRIEL, 2005)

A compreensão envolve uma rede complexa que se entrelaça com a visão, com o reconhecimento, com a memória, dedução, análise, conexão e avaliação no processo de leitura. É evidente que a leitura é por si só uma experiência que obriga a mente a estabelecer uma rede associativa tão intricada que chega a parecer ilimitada e imperscrutável, enredando dispositivos cognitivos, linguísticos, habilidades técnicas e envolvimento emocional. (GABRIEL, 2005)

Devido a toda essa complexidade, torna-se, então, indispensável esclarecer alguns pontos: o caráter processual e progressivo da compreensão que garante a possibilidade de passar a estruturas sempre mais profundas na abordagem do texto, enquanto as habilidades de base têm a tendência a constituir-se em meros automatismos. (OLMI, 2005)

Considerando a leitura como um ato cognitivo, propõe-se, com base em alguns autores, os seguintes passos de uma cadeia cognitiva para chegar-se à compreensão:

1. Movimento de apropriação – “é preciso rachar as palavras ou as frases para delas extrair os enunciados” (DELEUZE, 1998);

2. Movimento de interação – quando o leitor atribui novo significado ao texto, relacionando-o com a própria vida, ressignificando seu viver;

3. Movimento de experimentação – Deleuze (1998), aconselha a nunca interpretar um texto, mas sempre experimentá-lo, pois cognição é sempre algo vivido, experimentado;

4. Movimento de virtualização / atualização – Segundo Levy, “um pensamento se atualiza num texto e um texto numa leitura” (LÉVY, 1996), o impacto que um texto pode causar desencadeia no leitor o seu imaginário, configurando um novo mundo;

5. Movimento das emoções – Não há fenômeno cognitivo sem emoção. Segundo Maturana (1994, 1999), os seres humanos têm uma história de relações que constituem a organização do funcionamento da espécie e o amor aí é o elemento básico;

6. Movimento de invenção – quando se reinventa o texto com termos próprios;

7. Movimento da produção de diferença – apropriação do texto com a possibilidade de torná-lo outro, um texto singular;

8. Movimento da imaginação – quando são simuladas situações (PELLANDA, 2005).

Pelo exposto, ao constatar-se a leitura como um processo mental, difícil será dizer com precisão o que está se passando na mente do leitor. Para que esse conhecimento avance, pesquisadores aprofundaram seus estudos no processo de compreensão de textos. Os processos mentais são particulares a cada indivíduo e cada leitor produz diferentes compreensões de texto ao buscar e usar a informação.

COMPETÊNCIAS PARA BUSCAR E USAR A INFORMAÇÃO

Buscar e usar a informação constituem competências cruciais na sociedade da aprendizagem, envolve a busca ativa ou passiva da informação, planejamento, estratégias e motivação para atingir objetivos, monitoração de estratégias, conhecimento e definição de canais ou fontes de informações potenciais, competências para usar tecnologias da informação e avaliação desse processo.

O método mais utilizado para avaliar a informação eletrônica é aquele que se centra na satisfação do usuário e de suas necessidades de informação, e se baseia na análise e seleção de uma série de critérios que

devem estar presentes nos conteúdos proporcionados pelo mundo digital.

O uso da informação compõe-se de atividades em que o indivíduo se engaja para apreender a informação e transformá-la em conhecimento, compreendendo habilidades intelectuais como decodificação, interpretação, controle e organização do conhecimento. A decodificação e a interpretação por sua vez incluem atividades de leitura, estabelecimento de relações e conhecimento prévio e as novas informações, comparação de vários pontos de vista e avaliação. Controle e organização relacionam-se propriamente à organização da informação por meio de uso de instrumentos cognitivos, como recursos, esquemas, mapas conceituais e elaboração de textos.

De acordo com dados divulgados no documento ALFIN-EEES (2005), o primeiro passo para começar a avaliar um recurso eletrônico é identificar sua tipologia, isto porque na rede existe uma diversidade de recursos: diretório, motor de busca, revista eletrônica, portal, sítio comercial, sítio institucional, meio de comunicação, entre outros pontos. A aplicação de critérios e de indicadores de avaliação varia, portanto, em função das características de cada um deles, assim como do nível de profundidade com que se quer realizar a avaliação.

Os critérios de avaliação mais gerais, assim como os indicadores e parâmetros apresentados na análise de um site Web são os seguintes:

QUADRO 2 - CARACTERÍSTICAS PARA AVALIAÇÃO DA QUALIDADE DA INFORMAÇÃO. ALFIN.

CRITÉRIOS	INDICADORES
Autoria	cadastro do autor Informação sobre o autor Meio de contato (e-mail) Logotipo da organização Declaração de princípios e propósitos do site Web Avaliação externa
Atualização	Data de criação Data de atualização Informação atual e atualizada Existência de enlaces obsoletos ou incorretos
Conteúdo	Cobertura Exatidão, precisão e rigor Pertinência Objetividade

CRITÉRIOS	INDICADORES
Acessibilidade	Formato compatível com diferentes versões, navegador e resoluções de tela Versões alternativas de visualização Cumprimento normativo WAI Impressão correta Ajuda para a navegação e a compreensão de conteúdos Versões em outras línguas
Funcionalidade	Estrutura lógica: tabela, *menu* de conteúdos Pertinência e adequação dos títulos às seções Existência de mapa Webs com *links* Sistema de busca de conteúdos próprio
Navegabilidade	Menu de conteúdos Botões de navegação
Formato / desenho	Elegante, funcional e atrativo Combinação de cores, formas e imagens Tipografia textual adequada Homogeneidade de estilo e formato

Fonte: ALFIN/EEES (2005)

Diante da complexidade de buscar e usar a informação, impõe-se a necessidade e importância de estudos por parte de profissionais e usuários da informação. A competência, segundo Teixeira (2008), é a capacidade de mobilizar conhecimentos, valores e decisões para agir de modo pertinente numa determinada situação. As competências só podem ser constituídas na prática. Não é só o saber, mas o saber fazer. Aprende-se fazendo, numa situação que requeira esse fazer especializado. Em relação à habilidade, esta vem em decorrência da aquisição de competências.

Segundo Tobón (2004), **atitudes** são disposições afetivas à ação de maneira que as atitudes impulsionem o comportamento nos seres humanos. **Aptidão** se refere a potencialidades inatas que os seres humanos possuem e que necessitam ser desenvolvidas mediante educação. **Habilidades** implicam a formação em certas tarefas ou atividades para levá-las a cabo com perfeição; a pessoa habilidosa é capaz de realizar processos para os quais foi treinada com eficácia e eficiência.

Segundo Artigas e Tóbon (2006), as dimensões das competências estão assim sistematizadas: **cognitiva-declarativa –** assimilar noções, regras e princípios que regem o funcionamento de uma atividade laboral; **habilidades cognitivas –** processar as informações e gerar

conhecimento, necessitando da aprendizagem de estratégias de observação e análise da realidade circundante; **atuacional-procedimental** – refere-se a como se realizam as coisas e a como atuar na realidade; **social** – estabelecer relações com pessoas, para poder seguir o ritmo acelerado de mudanças, novidades e avanços; **comunicativa** – relaciona-se à cognição e às atividades de processamento da informação, subdivide-se em três competências básicas: interpretativa, argumentativa e propositiva; **ética** – relaciona-se aos benefícios para quem exerce uma profissão, supondo uma atividade que beneficia a outros.

Para Guy Le Boterf (2003), a competência consiste em saber mobilizar e combinar recursos. O profissional dispõe de uma dupla instrumentalização: a de recursos pessoais e a de recursos de seu meio. A primeira incorpora e constitui-se por saberes, saber-fazer, aptidões ou qualidades e por experiências acumuladas. Já a segunda é constituída por máquinas, instalações materiais, informações e redes relacionais.

A competência, segundo Teixeira (2008), é a capacidade de mobilizar conhecimentos, valores e decisões para agir de modo pertinente numa determinada situação. As competências só podem ser constituídas na prática. Não é só o saber, mas o saber fazer. Aprende-se fazendo, numa situação que requeira esse fazer especializado. Em relação à habilidade, esta vem em decorrência da aquisição de competências.

Vê-se, portanto, que, no contexto atual, ainda é indispensável o adequado manejo das tecnologias de informação e comunicação, o que vai requerer novas aprendizagens e estratégias de uso que facilitem o acesso à informação, fazendo-se frente a uma nova alfabetização, para adquirir dois tipos de competências:

a) **competência tecnológica** – ocupa-se da Teoria Prática do formato, do desenvolvimento, da seleção e utilização, da avaliação e gestão dos recursos tecnológicos utilizados na Educação e atende aos seguintes aspectos: conhecimentos científicos teóricos das TICs e meios de comunicação; habilidades de manejo; alfabetização audiovisual; alfabetização informática e telemática; valorização do impacto das TICs e dos meios de comunicação de massa na sociedade e na educação; conhecimento dos materiais disponíveis no mercado: meios de comunicação de massa, vídeos, *software,* espaços

Web e avaliação da qualidade técnica, pedagógica e funcional; conhecimento das possíveis aplicações em educação; planejamento, gestão e avaliação de atividades educativas com apoio tecnológico; delineamento de desenvolvimento de materiais educativos em suporte tecnológico; organização dos recursos pedagógicos centrais (ALFIN/EEES, 2005).

b) **competência informacional**, chamada também de educação em informação ou alfabetização informacional é um processo de aprendizagem, que se dá quando, a partir de um problema ou uma necessidade, buscam-se os recursos; localiza-se a informação; analisa-se, interpreta-se, sintetiza-se e comunica-se de maneira eficaz com outras pessoas; e, por último, avalia-se o resultado. Atuação que se centra em três áreas: **busca da informação –** habilidades de localizar e recuperar documentos; habilidades de manejar equipamentos tecnológicos; **Uso da informação –** habilidades de pensar; Habilidades de estudar e investigar; **Disseminação da informação –** habilidades de produzir e de representar (ALFIN/EEES, 2005).

Estas competências exigem um mínimo de habilidades para o pleno sucesso educativo no uso da Internet:

- Utilizar as principais ferramentas de Internet: navegadores, correio eletrônico, FTP, listas de distribuição e grupos de notícias, bate-papos, videoconferências, programas de navegação *off-line*... baixar informações, normas de estilo...
- Conhecer as características básicas de equipamentos e infra-estruturas informáticas necessários para acessar a Internet: ordenadores, *modens,* linhas telefônicas.
- Diagnosticar que informação se necessita
- Encontrar a informação que se busca e recuperá-la com agilidade. Para isso é preciso:
- Conhecer e usar os programas de busca, bibliotecas e bases de dados.
  – realizar com estes programas buscas por palavras e também buscas avançadas mediante combinações booleanas de descrições.
  – localizar listas de discussão, grupos de notícias, grupos de interesse na web relacionados com as temáticas que se procura.

- Avaliar a qualidade (autenticidade, atualidade) da informação que se queira. Há indicadores que se devem considerar: data da última atualização da página, o prestígio de seus autores ou patrocinadores, as ligações a outras páginas afins.
- Avaliar a idoneidade da informação obtida para ser utilizada em cada situação concreta. Não basta apenas encontrar a informação, há que se saber analisá-la e aplicá-la na resolução dos problemas que se apresentam.
- Aproveitar as possibilidades de comunicação que a Internet oferece (correio eletrônico, listas de discussão, grupos de notícias...) nas atividades laborais, culturais e recreativas.
- Avaliar a eficácia e eficiência da metodologia empregada na busca de informação e na comunicação através da Internet. Com esta revisão, se melhoram progressivamente as técnicas e estratégias empregadas e cada vez se atuará com mais eficácia e eficiência (ALFIN/EEES, 2005).

CONSIDERAÇÕES FINAIS

Neste artigo, procurou-se refletir sobre o movimento deliberado das tecnologias colaborativas e de acesso livre, no sentido de criar facilidades de acesso e uso da informação para o usuário, fazendo surgir meios mais flexíveis, atrativos e dinâmicos. No bojo desta proposta das TICs está implícita a necessidade do aprofundamento de estudos sobre o usuário e sua complexa manobra cognitiva; sobre o emaranhado de fontes de informação advindas do avanço tecnológico; sobre o desempenho da biblioteca na aplicação destas inovações científicas e tecnológicas; e sobre o papel mediador do bibliotecário neste contexto, de sofisticados mecanismos e ferramentas de busca em razão das múltiplas sintaxes. Neste quadro de mudanças, um grande desafio para o bibliotecário é enfrentar este período de transição, em que há usuários inseridos no mundo digital e outros não, principalmente nos países como o Brasil, onde a questão é mais complexa.

Pelo exposto, percebe-se que a as inovações no âmbito das TICs têm alcançado seus propósitos de criação de ferramentas cada vez

mais velozes, econômicas, interativas e hipertextuais, que aproximam o sujeito da rede, no entanto, numa sociedade desigual o problema não é apenas tecnológico, mas educacional e político, considerando que grande parte da população, como é sabido, não usufrui das ações de inclusão digital, tanto por falta de ferramentas tecnológicas como por carência de pré-requisitos cognitivos para a leitura das complexas linguagens tecnológicas e de conteúdo.

Não se pode negar o significado da Internet na intensificação da distribuição do conhecimento e do acesso à informação e na ampliação das possibilidades de educação do sujeito. No entanto, a complexidade e a multidimensionalidade desta rede é uma barreira a vencer.

A Web 2.0, o modelo construído com base numa filosofia de partilha e maior interatividade, apresenta dificuldades pela complexidade de seus processos informacionais, em função do uso de signos e linguagens híbridos, diversos da linguagem verbal escrita e impressa configurando um novo tipo de representação e organização do conhecimento. Embora as estatísticas sobre o uso da rede mostrem seu crescimento exponencial, as pesquisas sobre o perfil do usuário da rede, o que ele pesquisa, como pesquisa e para que pesquisa, destacam a forte relação sobre a educação e renda no acesso e uso da Internet. Quanto ao fator educacional, chama-se a atenção para o aspecto cognitivo no que se refere à necessidade de operações complexas e superiores para a codificação e decodificação dos conteúdos, que estão implícitas na cadeia de ações para a obtenção do conhecimento.

A este respeito, tem sido proposto um novo modelo de ensino que contempla as tecnologias da informação e da comunicação e os e-learning centrados na alfabetização informacional e digital. Esta iniciativa capacita a quem aprende a enfrentar criticamente os conteúdos e fazer-se autossuficiente, assumindo maior controle sobre seu próprio processo de aprendizagem. Este processo visa a desenvolver nos sujeitos uma ampla autonomia de pensamento que lhes permita adquirir, compartilhar e transferir conhecimento.

Nesta perspectiva de desenvolvimento de competências informacionais, a ALFIN/EEES apresenta em seu portal o movimento da Pedagogia Informacional: ensinar a aprender na Sociedade da Informação,

centrado na hipótese educativa ensinar a aprender, e, sobretudo, utilizar adequadamente a informação no processo ensino e aprendizagem, demandando uma nova arquitetura que aposta na aprendizagem para a vida (*lifelong learning*).

O propósito da Pedagogia Informacional é fazer com que docentes e estudantes assumam o papel de mediadores da informação, buscando reciprocidade, significado e transcendência desta informação, percorrendo o fluxo de um processo cognitivo nas dimensões de acesso, análise, interpretação, avaliação, produção etc., fechando o ciclo informacional e pedagógico – a informação como ponto de saída e de entrada. Vale a pena ressaltar neste processo, o valor de significação dado a determinado fenômeno; nesta perspectiva a informação para ser tratada requer conhecimento ou capacidade epistemológica, o que permite seu descobrimento e interpretação.

A pedagogia informacional objetiva estabelecer um verdadeiro diálogo entre o usuário-leitor e o ambiente informacional e construir aprendizagens significativas, por meio de metodologias inovadoras, tais como segmentos temáticos informacionais (*portfolios*), buscas especializadas, redes de trabalho temáticas, microbases de dados segundo assinaturas, leituras digitais alternativas, fóruns de discussão virtuais por meio de leituras digitais, concepções icônicas com imagens digitais, uso da imprensa digital, entre outras.

Deste modo, o modelo de biblioteca 2.0 teria função expressiva no contexto da Pedagogia informacioanal, no que diz respeito à aplicação das ferramentas da Web para a criação de serviços informacionais interativos, dinâmicos e partilhados e na capacitação de usuários na busca e uso da informação, constituindo-se numa rede social potencializadora da participação ativa dos usuários, a inteligência coletiva e o trabalho colaborativo, salvaguardando e fortalecendo a missão da instituição biblioteca e dos profissionais que nela atuam.

Para ajudar o vivenciar deste contexto, como está a formação do leitor?

O documento da ALFIN/EEES ressalta que desde Paulo Freire até nossos dias, o avanço e a inovação em matéria pedagógica foi relativamente lento, levando em consideração a dinamicidade científica em

outras áreas. A ALFIN/EEES reapresenta o conceito de alfabetização de Paulo Freire, para quem a alfabetização é "um aprender a escrever sua vida, como autor e como testemunha de sua história", ou seja, "alfabetizar é conscientizar", para comentar que na atualidade sucede o mesmo, pois tem-se a imperiosa necessidade de tomar consciência das coordenadas da sociedade do conhecimento, além das tarefas de criar, administrar, selecionar, processar e difundir a informação, como uma ferramenta fundamental para o desenvolvimento humano.

REFERÊNCIAS

ALCARÁ, Adriana Rosecler; CURTY, Renata Gonçalves. Blogs: dos diários egocentristas aos espaços de comunicação científica. In: TOMÁEL, Maria Inês (Org.). **Fontes de Informação na Internet**. Londrina: EDUEL, 2008. p. 81-96.

ALFIN/EEES. **Habilidades e competencias de gestión da información para aprender a aprender en el marco del espacio europeo de enseñanza superior**. Espanha : Universidade de Granada, 2005. Disponível em: www.um.es/dp-lengua-espa/.../habilidades-competencias.pdf Acesso em: 15 out. 2009.

ANDRADE, Laura Mariane de. **Aplicabilidade de recursos web 2.0 em bibliotecas**. 87 f. Trabalho de Conclusão de Curso (Graduação em Biblioteconomia) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, 2009. Disponível em: http://rabci.org/rabci/sites/default/files/TCC_LAURA1.pdf. Acesso em: 10 de abril 2010.

ARTIGAS, C. M. T.; Tobón, S. T. et al. (Coord). **El diseño del plan docente en información y documentación acorde con el espacio europeo de educación superior**: un enfoque por compentencias. Madrid: Facultad de Ciencias de la Documentación, Universidad Complutense de Madrid, 2006.

BAPTISTA, Ana Alice et al. Comunicação científica: o papel da open archives initiative no contexto do acesso livre. **Encontros Bibli**: revista eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, n. especial, p. 1-17, jan/jun., 2007. Disponível em: http://www.encontros-bibli.ufsc.br/bibesp/esp06/bibesp.esp.06.baptista20071.pdf Acesso em: 20 out. 2009

BASTOS, Flávia Maria; OLIVEIRA, Walter Clayton; VIDOTTI, Silvana Ap. B. Gregório. Ambientes informacionais: as bibliotecas digitais e os repositórios institucionais no contexto da web 2.0. In: CONGRESO DE LA CIBERSOCIEDAD, 4., 2009. **Crisis analógica, futuro digital**. Barcelona, ES: Observatorio para la Cibersociedad Citilab Cornellà, 2009. Disponóvel em: http://www.cibersociedad.net/congres2009/es/coms/ambientes-informacionais-as-

bibliotecas-digitais-e-os-repositorios-institucionais-no-contexto-da-web-20/1056/

BOTERF, Guy Le. **Desenvolvendo a competência dos profissionais**. Porto Alegre: Artmed, 2003.

CARVALHO, Maria Carmen Romcy de. Bibliotecas universitárias brasileiras e a implantação de repositórios institucionais. **Revista Informação e Universidade**, v.1, n. 0, 2009. Disponível em: http://www.siglinux.nce.ufrj.br/~gtbib/site/ 2009/06/implantacao-de-repositorios/. Acesso em: 22 mar 2010.

CASEY, Michael E. ; SAVASTINUK, Laura C. Library 2.0; service for the next-generation library. **Library journal**, sept. 2006. Disponível em http://www.libraryjournal.com/article/CA6365200.html Acesso em: 22 mar. 2010.

CHOO, Chun Wei. **A organização do conhecimento**: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões. 2. ed., São Paulo: Editora Senac, 2006.

DAHLBERG, I. Knowledge organization: is scope and possibilities. **Knowledge Organization**, Frankfurt, v. 20, n. 4, p. 211-222, 1993.

DELEUZE, Gilles; PARNET, Claire. **Diálogos**. São Paulo: Escuta, 1998.

FERREIRA, Sueli Mara Soares Pinto. Repositório institucional em comunicação: o projeto Reposcom implementado junto à Federação de Bibliotecas Digitais em Ciências da Comunicação. **Encontros Bibli**: revista eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, 2007. Disponível em: http://www.periodicos.ufsc.br/ index.php/eb/article/viewFile/459/453 Acesso em: 15 fev 2010.

FEURSTEIN, R.; FEURSTEIN, S. **A mediated learning experience (MLE)**: theorical, psychosocial and learning implications. London: Freund Publishing House, 1994.

FEUERSTEIN, R. **Instrumental enrichment**: an intervention program for cognitive modifiability. Baltimore: University Park Press, 1980.

GABRIEL, Rosangela. Compreensão em leitura: como avaliá-la? In: OLMI, Alba; PERKOSKI, Norberto (Org.). **Leitura e cognição**: uma abordagem transdisciplinar. Santa Cruz do Sul : EDUNISC, 2005.

KUHLTHAU, C. C. Inside the search process: information searching from the user's perspective. **Journal of the American Society for Information Science**, v. 42, n. 5, 1991.

KUHLTHAU, C. C. A principle if uncertainly form information seeking. **Journal of Documentation**, v. 49, n. 4, 1993.

KUHLTHAU, C. C. **Seeking Meaning:** process approach to Library and information services. Norwood: Ablex Publishing, 1993.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. 2.ed. rev. atual. Brasília, DF: Briquet de Lemos/Livros, 2004.

LÉVY, Pierre. **O que é o virtual?** São Paulo: Ed. 34, 1996.

MARUYAMA, M. Metaorganization of information. **Cibernética**, n. 4, 1965.

MATURANA, Humberto. **El sentido de lo humano**. Santiago: Dólmen, 1994.

MATURANA, Humberto. **Transformación en la convivência**. Santiago: Dólmen, 1999.

MENDONÇA, Ana Valéria M. **Os processos de comunicação e o modelo todos-todos**: uma relação possível com o Programa Saúde da Família. Brasília: Ed. do Departamento de Ciência da Informação e Documentação da Universidade de Brasília, 2007.

MEY, Eliane Serrão Alves. **Biblioteconomia envergonhada**. Set. 2009. Disponível em: http://www.ofaj.com.br/textos_conteudo.php?cod=264. Acesso em: 15 fev 2010.

MILANESI, L. **Biblioteca**. São Paulo: Ateliê, 2002.

MILLER, Paul. Web 2.0: building the new library. **Ariadne Issue 45**, Oct. 2005. Disponível em: http://www.ariadne.ac.uk/issue45/miller/intro.html Acesso em: 22 mar 2010.

MIRANDA, Antônio; LEITE, Cecília;Emir, SUAIDEN. A biblioteca híbrida na estratégia da inclusão digital na Biblioteca Nacional de Brasília. **Inclusão Social**, Brasília, v. 3, n.1, p. 17-23 out. 2007/mar. 2008. Disponível em: http://antoniomiranda.com.br/ciencia_informacao/biblioteca_hibrida.html Acesso em: 22 mar. 2010

MORIN, Edgar. **A cabeça bem-feita**: repensar a reforma, reformar o pensamento. 15. ed. Rio de Janeiro : Bertrand Brasil, 2008.

MORIN, Edgar. **Ciência com consciência**. 2. ed. rev. Rio de Janeiro: Bertrand Brasil, 1998.

MORIN, Edgar. **Los siete saberes necesarios para la educación del futuro**. Barcelona : Paidós Ibérica, 2002.

MURLEY, Diane. What is all the fuss about library 2.0? **Law library journal**, v. 100, n. 1, 2008.

NEUROTH, Heikel. A biblioteca digital e o conceito de uma nova cultura informacional. Palestra proferida no Seminário **A sociedade do conhecimento e a biblioteca digital**, Salvador-BA: Instituto Goethe; Fundação Pedro Calmon, 2009. Disponível em: http://www.twiki.ufba.br/twiki/pub/ICI/NoticiasIci/Neuroth_port1.pdf. Acesso em: 30 mar. 2010.

OLMI, Alba. PERKOSKI, Norberto (Org.). **Leitura e cognição**: uma abordagem transdisciplinar. Santa Cruz do Sul : EDUNISC, 2005. 256p.

OLMI, Alba. Leitura, literatura e ciências cognitivas: uma aliança difícil mas necessária. In: OLMI, Alba; PERKOSKI, Norberto (Org). **Leitura e cognição**: uma abordagem transdisciplinar. Santa Cruz do Sul : EDUNISC, 2005.

PALCHEVICH, Diana Rosa Rodriguez. Integrando la web social 2.0 a las bibliotecas populares pampeanas para un desarrollo regional más allá de las distancias y desiertos. In: CONGRESO MUNDIAL DE INFORMACIÓN Y CONOCIMIENTO: aspectos tecnológicos, 2., Santo Domingo, Rep. Dominicana, 2009. Fundación Ciencias de la Documentación.

PELLANDA, Nize Maria Campos. Leitura como processo cognitivo complexo. In: OLMI, Alba. PERKOSKI, Norberto (Org). **Leitura e cognição**: uma abordagem transdisciplinar. Santa Cruz do Sul : EDUNISC, 2005. 256p.

SARACEVIC, Tefko. Ciência da informação: origem, evolução e relações. **Perspec. Ci. Inf.**, Belo Horizonte, **v. 1**, n. 1, p. 41-62, jan./jun. 1996

SARACEVIC, Tefko. I. Relevance: a review of and a framework for the thinking on the notion in information, science. In: BRAGA, Gilda Maria; PINHEIRO, Lena Vânia Ribeiro (Org.). **Desafios do impresso ao digital**: questões contemporâneas de informação e conhecimento Brasília: IBICT: Unesco, 2009. p. 15-69.

SARACEVIC, Tefko. 2. Relevance: a review of the literature and a framework for thinking on the notion in information, science. Part II: nature and manifestations of relevance. In: BRAGA, Gilda Maria; PINHEIRO, Lena Vânia Ribeiro (Org.). **Desafios do impresso ao digital**: questões contemporâneas de informação e conhecimento Brasília: IBICT: Unesco, 2009. p. 71-116.

SARACEVIC, Tefko. 3. Relevance: a review of the literature and a framework for thinking on the notion in information science. Part III: behavior and effects of relevance. In: BRAGA, Gilda Maria; PINHEIRO, Lena Vânia Ribeiro (Org.). **Desafios do impresso ao digital**: questões contemporâneas de informação e conhecimento Brasília: IBICT: Unesco, 2009. p. 117-168.

SOUZA, Ana Maria Martins de; DESPRESBITERIS, Lea; MACHADO, Osny Telles Marcondes. **A mediação como princípio educacional** – bases teóricas de Reuven Feuerstein. São Paulo: Editora Senac São Paulo, 2004.
STERNBERG, Robert J. **Psicologia cognitiva**. Porto Alegre: Artes Médicas,2000.

TEIXEIRA,G. **Significado de competência. Ensino e aprendizagem. Ser professor universitário**. Disponível em: http://www.serprofessoruniversitario.pro.br/ler.php?modulo=12&texto=728 Acesso em: 28 set. 2009

TOBÓN,S . **Formación basada en competencias. Pensamiento complejo. Diseño curricular Y Didática**. Bogotá: Ecoe, 2004.

VYGOTSKY, L.S. **Pensamento e linguagem**. São Paulo: Martins Fontes, 1993

WEITZEL, Simone da Rocha. **Os repositórios de e-prints como nova forma de organização da produção científica**: o caso da área das ciências da comunicação no Brasil. São Paulo, 2006. 361f. Tese. (Doutorado em Ciência da Informação) – Universidade de São Paulo.

Insumos conceituais e práticos para iniciativas de repositórios institucionais de acesso aberto à informação científica em bibliotecas de pesquisa

Sely Maria de Souza Costa
Fernando César Lima Leite

INTRODUÇÃO

Repositório institucional foi o termo utilizado para representar um novo serviço bibliotecário cuja ênfase é constituir, gerenciar e, sobretudo, disseminar amplamente coleções digitais de informação científica, de modo que a comunicação, acesso e uso de resultados de pesquisa fossem expandidos. Crow (2002a), ao definir repositório institucional como um arquivo digital de produtos intelectuais criados por docentes, pesquisadores e estudantes de uma instituição, e acessíveis com poucas ou nenhuma barreira, sugeriu também atributos que explicitam com propriedade os requisitos e a natureza de um repositório institucional. Trata-se de quatro atributos, nomeadamente: institucionalmente definido; científica ou academicamente orientado; cumulativo e perpétuo; aberto e interoperável. São, de fato, atributos que parecem ter alcançado o consenso da comunidade internacional, pois traduzem a capacidade própria dos repositórios institucionais constituírem, ao mesmo tempo e de modo complementar, duas funções. A primeira constitui uma nova e mais adequada alternativa de gestão da informação científica, enquanto que a segunda se traduz em componente do atual e complexo sistema de comunicação científica. Ambas as

funções refletem com clareza transformações no comportamento e necessidades informacionais de usuários de informação científica e as demandas institucionais de lidar apropriadamente com o volume crescente de informação científica digital.

Sobre ser "institucionalmente definido", é importante pontuar recente debate realizado por meio do principal fórum de discussão do tema, *American Scientist Open Access Forum*[1], em que se observou que os repositórios digitais podem ser classificados em diferentes tipos, dos quais se destacam os institucionais e os disciplinares ou temáticos. Além de seus limites serem definidos pelas fronteiras da instituição, os repositórios institucionais devem ser assim considerados se satisfizerem a pelo menos duas condições. A primeira é que sejam oficialmente reconhecidos pela instituição, por meio da implementação de políticas de depósito compulsório e outras que garantam sua existência. A segunda é que seus conteúdos cubram a maior parte das áreas de ensino e pesquisa da instituição, como demonstração de seu reconhecimento e aceitação.

No que concerne a ser "científica ou academicamente orientado", é importante ressaltar duas questões, as quais têm relação com decisões presentes nas políticas de funcionamento. A primeira é que "cientificamente orientado", deve significar conteúdo reconhecido e validado pela comunidade científica, o que, por sua vez, aponta, necessariamente, para as "publicações" científicas, especialmente o artigo de periódico. A segunda é que "academicamente orientado" flexibiliza a perspectiva da validação alcançada pelo *peer review*, e contempla também outras formas de comunicação científica, mais informais, inclusive material de ensino.

Em relação a ser "cumulativo e perpétuo", é importante notar que essa é uma das motivações iniciais da proposta de repositórios institucionais, qual seja, preservar a produção intelectual de uma instituição para lhe garantir acesso amplo e irrestrito. Isso, por sua vez, requer a implementação de políticas de depósito, de preservação, de acesso e de uso.

---

1 http://amsci-forum.amsci.org/archives/American-Scientist-Open-Access-Forum.html

Por fim, ser "aberto e interoperável" é uma das principais características responsáveis pelo ressonante aumento do impacto dos resultados de pesquisa e visibilidade da produção, do pesquisador e da instituição. Tal capacidade resulta de padrões tecnológicos adotados que permitem à comunidade científica descobrir e recuperar informações armazenadas em repositórios institucionais a partir da exposição e coleta automatizada de metadados por variados mecanismos de buscas. Aliadas a esse potencial tecnológico estão as políticas institucionais de acesso aberto, que, além de estimularem o depósito de conteúdos no repositório, devem harmonizar aspectos relacionados com direitos autorais e as necessidades de acesso e uso a eles.

São essas as questões, portanto, que norteiam a discussão por todo este capítulo, que se compõe de quatro seções cujo objetivo é discutir aspectos teóricos e práticos relacionados com a criação e funcionamento de repositórios institucionais em bibliotecas de pesquisa.

REPOSITÓRIOS INSTITUCIONAIS: CONCEITOS E NATUREZA

Antes da definição do que constitui um repositório institucional de acesso aberto à informação científica, é importante tecermos considerações acerca de um conceito mais amplo: repositórios digitais. A expressão repositórios digitais, no contexto do movimento mundial em favor do acesso aberto, é utilizada para denominar os vários tipos de provedores de dados que constituem vias alternativas de comunicação científica. Cada um dos tipos de repositórios digitais possui funções específicas no sistema de comunicação científica e aplicações próprias voltadas para o ambiente no qual será utilizado. Com base no Digital Repositories Infrastructure Vision for European Research – DRIVER[2] e, especialmente, em estudos por ele financiados (WEENINK, et al. 2008; SWAN, 2008; VAN WEIJNDHOVEN, K.; VAN DER GRAAF, 2007), considera-se que, de maneira geral, os repositórios digitais podem ser de três tipos:

2 http://www.driver-repository.eu/

- repositórios disciplinares ou temáticos: voltados a comunidades científicas específicas. Tratam, portanto, da produção intelectual de áreas do conhecimento em particular. Exemplo: EPrints in Library and Information Science (E-LIS)[3] e arXiv.org[4];
- repositórios de teses e dissertações (Electronic Theses and Dissertation – ETDs): repositórios que lidam exclusivamente com teses e dissertações. Muitas vezes a coleta das muitas ETDs é centralizada por um agregador. Exemplo: Biblioteca Digital de Teses e Dissertações da Universidade de Brasília (BDTD/UnB)[5] e BDTD[6];
- repositórios institucionais: dedicados à produção intelectual de uma instituição, especialmente universidades e institutos de pesquisa. Exemplo: e-Prints Soton - repositório de Pesquisa da Universidade de Southampton[7];

Sabe-se que o emprego do termo repositório não é novo, especialmente na informática. Entretanto, os conceitos sobre os quais os repositórios digitais se desenvolvem e as funções às quais são destinados constituem a verdadeira inovação no contexto específico da comunicação na ciência. Por esta razão, ao se falar em repositório institucional de acesso aberto à informação científica, compreendem-se, necessariamente, os atributos propostos por Crow, mencionados anteriormente e que se confirmaram por completo quando da proposição do DSpace, no âmbito do Massachusetts Institute of Technology (MIT), como apresentado por Barton (2003).

Com base em tais atributos, sustenta-se que todo repositório institucional de acesso aberto pode ser considerado um tipo de biblioteca digital. Entretanto, contrariamente, nem toda biblioteca digital pode ser considerada um repositório institucional. Embora não haja na

---

3 http://eprints.rclis.org/

4 http://arxiv.org/

5 http://bdtd.bce.unb.br/tedesimplificado/

6 http://bdtd.ibict.br/

7 http://eprints.soton.ac.uk/

literatura discussão conceitual mais aprofundada acerca das diferenças ou similaridades entre repositórios institucionais e bibliotecas digitais, assume-se que, no contexto do acesso aberto e da comunicação científica, há, pelo menos, cinco diferenças entre os dois tipos serviços.

A primeira, como expresso anteriormente, é que repositórios institucionais de acesso aberto à informação científica lidam exclusivamente com a produção intelectual de uma instituição. Adicionalmente, dois aspectos derivados dessa questão se complementam, ao mesmo tempo em que se interpõem. De um lado, a fronteira institucional impõe restrição à aquisição e armazenamento de conteúdos cuja totalidade de autores não pertence à instituição. De outro, a natureza científica e acadêmica do repositório adverte contra conteúdos que possuem outra finalidade (por exemplo: documentos administrativos), como pode ser o caso de bibliotecas digitais.

A segunda diferença refere-se ao autoarquivamento – ou autodepósito, isto é, o depósito de conteúdos pelos dos próprios autores ou por um mediador – e à interoperabilidade, os quais constituem características que devem estar presentes em um repositório institucional, mas não necessariamente em uma biblioteca digital.

A terceira diferença diz respeito à maneira como *software* de repositórios institucionais são desenhados. Todas as funcionalidades e recursos das plataformas tecnológicas para a construção de repositórios institucionais pautam-se nas peculiaridades que envolvem os processos de gestão da informação científica e, sobretudo, nas características dos processos de comunicação científica. Bibliotecas digitais, entretanto, não necessariamente devem estar ligadas a esse contexto.

Portanto, para serem considerados como tal e, de fato, comporem o novo sistema de comunicação científica, as características mencionadas nos três parágrafos anteriores devem estar necessariamente presentes em um repositório institucional. Não necessitam, no entanto, estar obrigatoriamente presentes em uma biblioteca digital para que esta possa ser considerada como tal.

Um repositório institucional de acesso aberto constitui, desse modo, um serviço de informação científica – em ambiente digital e interoperável – dedicado ao gerenciamento da produção científica

e/ou acadêmica de uma instituição (universidades ou institutos de pesquisa). Contempla, por conseguinte, a reunião, armazenamento, organização, preservação, recuperação e, sobretudo, a ampla disseminação da informação científica produzida na instituição.

ELIMINANDO MAL ENTENDIDOS ACERCA DE REPOSITÓRIOS INSTITUCIONAIS

Tanto na literatura especializada, particularmente a nacional, quanto na prática profissional, percebe-se uma série de equívocos e distorções relacionadas com repositórios institucionais de acesso aberto e suas funções no sistema de comunicação científica. Visando a esclarecer parte desses mal entendidos, discorre-se a seguir sobre alguns equívocos e distorções.

a. *Repositórios institucionais X periódicos científicos*. Repositórios institucionais são diferentes e não concorrem com periódicos científicos. Registro (estabelecimento da propriedade intelectual), certificação (validação e qualidade dos resultados de pesquisa), circulação (assegurar a acessibilidade aos resultados de pesquisa) e arquivamento (preservação para uso futuro) constituem funções básicas do periódico científico (ROOSENDAAL; GEURTS, 1998). Com exceção da certificação, atribuição exclusiva dos periódicos, todas as outras funções são pontecializadas pelos repositórios institucionais, que constituem na realidade vias alternativas de comunicação científica. Por esta razão, repositórios institucionais não concorrem com periódicos, mas, sim, potencializam parte de suas funções.

b. *Repositórios institucionais e peer review (avaliação pelo pares).* Nas discussões iniciais sobre repositórios institucionais aventou-se a ideia da incoporação do processo de avaliação por pares. Tal possibilidade não mostrou qualquer sinal de aceitação por parte da comunidade científica e, muito rapidamente, deixou de fazer parte das estratégias do movimento em favor do acesso aberto. A avaliação por pares é uma exclusividade de periódicos e outros veículos de publicação científica e de modo algum é viabilizada

ou deve ser almejada por repositórios institucionais.

c. *Repositórios institucionais como veículos de publicação científica.* Um equívoco terminológico recorrentemente cometido por gestores de repositórios institucionais e também por pesquisadores do tópico contribui para a fragilização dos argumentos do acesso aberto. Quando expressões como "artigos científicos **publicados** no repositório", e suas variações, são utilizadas nos discursos de convencimento ou sensibilização, tanto pesquisadores quanto gestores acadêmicos ficam confusos e não compreendem as diferenças entre o repositório institucional e os veículos de publicação científica que comumente utilizam. É importante frisar que repositórios institucionais são veículos de maximização da **disseminação** de resultados de pesquisa, dado que tornam seus conteúdos disponíveis e acessíveis amplamente. Do ponto de vista formal, entretanto, a **publicação** é uma prerrogativa de periódicos e outros veículos. Esse mal entendido reforça a falsa ideia de concorrência entre repositórios e periódicos. Como o contexto é o mesmo, a comunicação científica, é fundamental o uso apropriado das expressões, sobretudo por tratar-se de esforços de adoção de estratégias de acesso aberto.

d. *Repositórios institucionais ≠ de sistemas de gerenciamento de bibliotecas.* Diferentemente de sistemas de gerenciamento de bibliotecas, cuja ênfase está na manutenção, desenvolvimento e controle do acervo, repositórios institucionais lidam exclusivamente com a produção científica e acadêmica da instituição, em formato digital, com vistas à sua ampla disseminação e visibilidade. Por outro lado, a fim de potencializar o uso e benefícios na instituição, ainda na fase de planejamento e implementação de repositórios institucionais, deve-se considerar a possibilidade de integração com processos e sistemas com funções próximas já existentes na instituição, como é o caso dos sistemas de gerenciamento de bibliotecas. No Brasil, a Empresa Brasileira de Pesquisa Agropecuária (Embrapa), por exemplo, possui uma bem sucedida integração entre o seu repositório institucional e o sistema de gerenciamento de suas bibliotecas.

e. *Repositórios institucionais e o problema da proteção do conhecimento.* Um dos questionamentos equivocados mais comuns enfrentados por gestores de repositórios institucionais, proveniente, na maioria das vezes, de pesquisadores, é a consideração de que o depósito da produção científica contribui para que os resultados de pesquisa sejam roubados e apropriados por outros pesquisadores, instituições ou países. De fato, o repositório institucional promove maior visibilidade e uso da informação científica. Contudo, a preocupação com a proteção de conhecimentos sensíveis ou resultados de pesquisa pantenteáveis deve existir antes da sua publicação (ou aceitação para publicação), que, na realidade, ocorre antes do depósito em repositório institucional. A publicação de resultados de pesquisa em periódicos científicos ou outro veículo de comunicação *per se* já garante, de algum modo, que interessados tenham acesso. As políticas de depósito obrigatório, imprescindíveis para o povoamento dos repositórios, geralmente têm como objeto a informação científica publicada formalmente. O depósito de outros tipos de informação científica não constitui, de modo geral, a obrigatoriedade das políticas institucionais.

f. *Repositórios institucionais aumentam a possibilidade de plágio.* Uma questão complementar à da proteção ao conhecimento é a de que, de fato, por tornar a informação científica amplamente disponível, repositórios aumentam, de certo modo, a possibilidade de plágio. No entanto, ao mesmo tempo, aumentam exponencialmente o registro da autoria pois, quanto mais disponível e acessível um trabalho, mais sua propriedade intelectual é reforçada. Um outro aspecto a mencionar é que a disponibilidade e acessibilidade expandidas aumentam também a descoberta do plágio. É muito mais complexa a identificação de uma obra plageada quando esta se encontra apenas em formato impresso na estante da biblioteca. Por outro lado, quando a obra plageada encontra-se acessível em meio digital *online*, como é o caso dos repositórios institucionais, a identificação do plágio é facilitada, seja manualmente, seja por meio do uso de *software* para identificação de plágios.

g. *Repositórios institucionais = informação científica ou academicamente orientada*. Como mencionado anteriormente, para serem considerados como tal, repositórios institucionais de acesso aberto devem reter alguns atributos. Um dos mais relevantes é que a finalidade da informação a ser gerenciada seja fundamentar o avanço científico e tecnológico, em atividades de pesquisa e ensino, e que o público alvo seja a comunidade científica e acadêmica. Com isso, qualquer variação na natureza da informação típica de repositórios institucionais é suficiente para desqualificá-lo como tal, tornando-se um outro serviço de informação, e não mais um repositório institucional de acesso aberto.
h. *Repositórios institucionais e seus usuários*. Mckay (2007) divide os usuários de repositórios institucionais em três principais grupos: autores, usuários finais (leitores) e criadores/mantenedores de dados (aqueles criam metadados, carregam arquivos e geralmente zelam pelas coleções do repositório). Segundo autor, dos três segmentos, os usuários finais são o que menos se tem estudado e, por consequência, o segmento menos conhecido. Diferentemente das coleções físicas de uma biblioteca, as coleções digitais possuem usuários potenciais em todo o mundo. Portanto, um repositório institucional deve ser planejado, implementado e permanecer em funcionamento tendo em vista que sua audiência é tanto local quanto global. Desse modo, todas as decisões devem ser tomadas considerando que os objetivos e preocupações de seus usuários variam em função do papel que desempenham. No que tange aos usuários finais é importante ter em mente que estão dispersos no mundo, e não apenas localmente na instituição.
i. *Repositórios institucionais e memória da instituição*. Não resta dúvida de que a plena adoção e funcionamento de um repositório institucional contribui adequamente para a composição do acervo da memória institucional. Entretanto, é importante não perder de vista sua finalidade primordial, que é aumentar a visibilidade dos resultados de pesquisa, do pesquisador e da instituição por meio de processos adequados de gestão da informação científica. Se tal finalidade for plenamente alcançada muito provavelmente

sua função de memória institucional também o será. O contrário poderá não resultar no mesmo.

j. *Repositórios institucionais e o foco na tecnologia*. A instalação e configuração de um *software* de repositório institucional não garantem sua existência como tal. Como será visto mais adiante, há providências que antecedem e que vêm após o *software*. Um outro equívoco relacionado com este aspecto é o fato de que muitas instituições não acadêmicas, ao adotarem um *software* (como o DSpace), geralmente para a construção de bibliotecas digitais, consideram que estão criando um repositório institucional. Um software de modo algum é suficiente para determinar a existência de um repositório institucional, mas sim os atributos que serão impressos no serviço e a função que este exercerá na instituição em relação ao sistema de comunicação científica. Por esta razão, instituições não acadêmicas ou não científicas podem e serão beneficiadas, caso adotem, por exemplo, o DSpace em seus projetos de bibliotecas digitais. Entretanto, certamente não estão criando um repositório institucional, tal como compreendido no contexto da comunicação científica e do acesso aberto, berços dos repositórios institucionais.

ABORDAGENS PARA CRIAÇÃO E IMPLEMENTAÇÃO DE REPOSITÓRIOS INSTITUCIONAIS

Mais do que um serviço, repositórios institucionais devem ser compreendidos como manifestação da reestruturação do sistema de comunicação científica. Sua emergência representa a materialização de uma filosofia de acesso aberto que se instaura na comunidade científica mundial. Constituem, desse modo, poderosa alternativa que, do ponto de vista da disponibilidade e acesso irrestrito à informação, potencializa a produção do conhecimento. Já do ponto de vista da disseminação da informação, proporciona a visibilidade e maximização do impacto de resultados de pesquisa, por meio da ampliação do seu acesso.

É importante salientar que, embora os benefícios decorrentes da

adoção de repositórios institucionais ocorram a partir de possibilidades tecnológicas, a tecnologia em um projeto de repositório institucional deve ser considerada como um dos diversos elementos que o integram. Isso significa que a instalação de *software* de repositório institucional não leva, necessariamente, ao sucesso da iniciativa. Contrariamente, o foco na tecnologia decerto conduz a iniciativa ao insucesso. É fundamental que um repositório institucional nasça com propósitos bem definidos, a partir de um planejamento elaborado e devidamente contextualizado. Ou seja, um repositório institucional deve surgir com funções a serem desempenhadas tanto internamente, na instituição, quanto no complexo sistema de comunicação científica global.

ORIENTAÇÃO PARA A CRIAÇÃO DE REPOSITÓRIOS: ABORDAGEM RÍGIDA OU FLEXÍVEL?

No que diz respeito à sua orientação, iniciativas de repositórios institucionais em todo o mundo podem ser agrupadas em duas abordagens principais, as quais convêm denominar rígida e flexível, respectivamente. É importante que os responsáveis pelo seu planejamento e construção conheçam e estejam cientes das implicações da escolha de uma ou outra abordagem.

Na abordagem rígida estão iniciativas – como o *Glasgow ePRINTS Service*[8] – cuja orientação é de que os repositórios institucionais devem priorizar conteúdos que foram submetidos ao processo de avaliação pelos pares, especialmente artigos de periódicos, sejam eles pré-prints ou pós-prints . Nesse caso, o principal argumento diz respeito ao controle de qualidade conferido pela avaliação por pares às informações que o repositório gerencia. Tal argumento contribui, indiscutivelmente, para sua credibilidade e adoção por parte da comunidade, como alternativa de comunicação científica.

A abordagem rígida preconiza que repositórios institucionais devem responder, preponderantemente, à função de potencializar a comunicação científica formal, visando especialmente à maximização

8 http://eprints.gla.ac.uk/

dos impactos dos resultados das pesquisas por meio da disseminação ampla e irrestrita de artigos de periódicos científicos, principal argumento do movimento de acesso aberto à informação científica. Qualquer variação desse foco significa enfraquecimento dos argumentos do movimento em prol do acesso aberto, que tem por missão principal aumentar o impacto da pesquisa por meio da maximização do acesso. Por essa razão, sob essa abordagem, os repositórios institucionais devem priorizar a literatura científica avaliada pelos pares, o que, essencialmente, está ligado ao sistema de publicações científicas.

Contudo, a abordagem flexível para repositórios institucionais, caso seja comparada com a rígida, amplia a sua destinação e contempla, além da literatura científica avaliada por pares, outros conteúdos de natureza acadêmico-científica produzidos por membros da instituição. Um exemplo é o DSpace at Cambridge[9]. Isso não significa que a abordagem flexível contrapõe-se à rígida, mas sim que a flexível agrega outros elementos à rígida. Ou seja, também são consideradas outras formas de conteúdo e estruturas de comunicação científica, tais como a literatura cinzenta, dados brutos de pesquisa, simulações, imagens, vídeos, relatórios de pesquisa, objetos de aprendizagem, entre outros.

Nessa perspectiva, um repositório institucional conjuga aspectos da comunicação científica formal e informal. Um dos argumentos dessa abordagem reforça que o conhecimento científico não é produzido exclusivamente a partir daquilo que já foi avaliado e é publicado formalmente, mas também daquilo que é veiculado informalmente, tornando o processo de comunicação científica mais flexível. Por essa razão, gestores de repositórios institucionais podem trabalhar para que o resultado dos esforços da gestão dos processos de identificação, armazenamento, preservação, recuperação e disseminação ampla da produção intelectual da universidade (seja ela avaliada pelos pares ou não) proporcione tanto a visibilidade do pesquisador e da instituição, quanto a promoção de condições férteis para a produção de novos conhecimentos. É importante frisar, no entanto, que essa flexibilidade não quer dizer que tudo aquilo que é produzido nos limites da universidade poderá ser incluído no repositório.

---

9 http://www.dspace.cam.ac.uk/

Sob a orientação flexível, portanto, repositórios institucionais devem compreender e atender também a demandas específicas de gestão institucional (apoiando tanto a produção quanto a comunicação do conhecimento) que parecem não ser adequadamente suportadas pelo sistema de comunicação científica formal. Dessa maneira, as preocupações da perspectiva flexível estão relacionadas com o sistema de comunicação científica como um todo, inclusive o sistema de publicações científicas, sem, entretanto, abranger aspectos da comunicação organizacional.

Por fim, em nível macro, a abordagem rígida volta-se exclusivamente para o sistema de publicações científicas. Este pode ser entendido como um subsistema constituinte de um complexo maior e abrangente, formado por outros elementos, denominado sistema de comunicação científica, cujos limites, incluindo o sistema de publicações, representam o interesse da abordagem flexível. Assim, a relação entre as abordagens e a jurisdição do sistema de comunicação científica como um todo pode ser ilustrada como na figura 1.

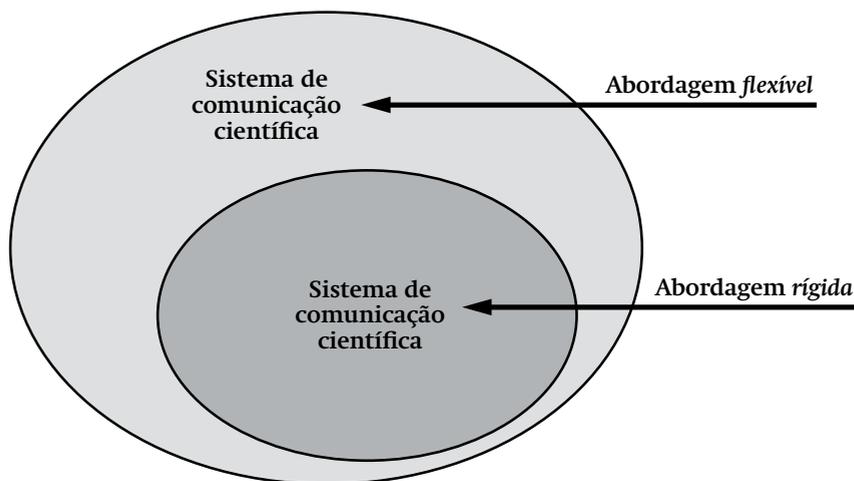



Figura 1 - Sistema de comunicação e de publicações científicas: abordagens rígida e flexível.
Fonte: Leite (2009).

A decisão por uma ou outra abordagem deverá ser fundamentada nas necessidades da instituição e objetivos estabelecidos para o repositório. As vantagens, desvantagens e recomendações para a adoção de uma ou outra abordagens são detalhadamente descritas em Leite (2009).

DIRETRIZES PARA A CONSTRUÇÃO DE REPOSITÓRIOS INSTITUCIONAIS

Para que repositórios institucionais exerçam satisfatoriamente funções e papéis preconizados pelo movimento de acesso aberto, é fundamental que a equipe responsável por sua construção possuam a dimensão das implicações contextuais, teóricas e práticas que envolvem o seu planejamento, implementação e funcionamento no âmbito de bibliotecas de universidades e institutos de pesquisa.

A maior parte das iniciativas de criação de repositórios institucionais é parte ou é realizada nas bibliotecas das instituições de ensino e pesquisa. Isso certamente ocorre por que os processos envolvidos nas rotinas de um repositório institucional possuem natureza muito próxima e similar aos trabalhos desenvolvidos em ambientes digitais por bibliotecas e bibliotecários.

Leite (2009) propôs recomendações para a construção de repositórios institucionais de acesso aberto à informação científica que na realidade constituem um conjunto de instruções que sistematizam aspectos que devem ser considerados no momento da elaboração e execução de um projeto de repositório institucional em universidades e institutos de pesquisa. O conjunto de instruções propostas pelo autor, resumidas na figura 2, fundamenta-se em melhores práticas de experiências bem-sucedidas em iniciativas de reconhecida importância, bem como em recomendações de manuais elaborados por instituições de renome, como o MIT, relatos de experiência e resultados de pesquisas de reconhecida importância na comunidade internacional que se dedica ao estudo do tópico (BARTON; WATERS, 2004; CROW, 2002a, 2002b; LYNCH, 2003; JONES et al., 2006; SWAN, 2008; SPARC, 2008; PROUDMAN, 2008 e outros).

O autor propõe três grandes fases interdependentes constituídas de atividades que devem ser cumpridas a fim de que a iniciativa de construção do repositório institucional seja bem sucedida. Evidentemende, esta não é a única maneira de se construir repositórios institucionais, mas o esquema proposto cobre aspectos relevantes que devem ser considerados nesse tipo de empreendimento.

Figura 2 - Fases da construção de repositórios institucionais de acesso aberto
Fonte: Leite (2009).

Planejar, implementar, assegurar a participação da comunidade e adotar estratégias para constituir o sistema global e aberto de comunicação são as grandes fases propostas. O detalhamento de cada uma das fases, atividades e todas as recomendações encontram-se descritos em Leite (2009).

Planejamento, implementação, participação da comunidade e adoção de estratégias têm sido, de fato, fundamentados em ideias que, bem formuladas, contribuem para o sucesso das iniciativas. Algumas iniciativas bem-sucedidas, em todo o mundo, estão citadas nos parágrafos anteriores. É importante, contudo, destacar aspectos que, observados em experiências de várias instituições ao redor do mundo, têm norteado a criação de repositórios nas universidades brasileiras.

Costa (2008) tem chamado a atenção para dois modos de operação, nomeadamente **sensibilização** e **ação real**, nas abordagens necessárias ao trabalho com repositórios digitais de acesso aberto. Isso porque a adoção desses dois modos de operação deve orientar as iniciativas com o acesso aberto, de modo geral, e com repositórios de acesso aberto (via verde), em particular, e periódicos científicos de acesso aberto (via dourada).

Sensibilização corresponde à adoção de abordagem *top-down*, que compreende o trabalho dos principais *stakeholders* do acesso aberto, envolvendo, basicamente, a conscientização de gestores de universidades, formuladores de políticas do governo e decisores de agências de fomento. Como se pode notar, são abordagens que representam o trabalho de sensibilização para obtenção do apoio necessário e do cometimento desses executivos da comunidade científica com relação às questões do acesso aberto. Tal sensibilização é especialmente necessária tendo em vista as reações iniciais de pesquisadores ao movimento do acesso aberto, reações que, em muitos casos, resultam da falta de interesse ou do medo de interpretações errôneas em relação ao autoarquivamento e à qualidade dos periódicos científicos de acesso aberto. Portanto, se decisores se tornam cometidos e comprometidos com o acesso aberto, é altamente provável que pesquisadores respondam mais positivamente ao movimento e às iniciativas a ele relacionadas.

Ação real corresponde à adoção da abordagem *bottom-up*, que tem se mostrado altamente recomendável, por mostrar resultados mais cedo, sendo, por conseguinte, fator positivo no convencimento da

comunidade. A abordagem *bottom-up* requer a elaboração de políticas por parte dos atores diretamente responsáveis pelas iniciativas, assim como a própria implementação destas. Stevan Harnad – autodenominado "arquievangelista" do acesso aberto –, em um sem número de interpelações feitas por meio de listas de discussão e mesmo de publicações e palestras em eventos, tem reforçado a necessidade de elaboração de políticas de depósito compulsório para repositórios institucionais. Mais que isso, a elaboração de políticas de funcionamento desses repositórios e de todas as demais políticas necessárias para seu pleno e bem sucedido funcionamento é atribuição dos atores responsáveis por criá-los e mantê-los. Ou seja: bibliotecários. As ações orientadas por essas políticas envolvem, portanto, o trabalho conjunto de bibliotecários e pesquisadores, compreendendo, pelo menos, três tipos. O primeiro diz respeito a visitas dos bibliotecários às diversas unidades acadêmicas da universidade visando à apresentação da iniciativa. O segundo relaciona-se a participações em reuniões de colegiados, com possíveis demonstrações dos avanços obtidos. Finalmente, deve haver programas de treinamento para o depósito no repositório. A principal ação real, no entanto, consiste no povoamento propriamente dito do repositório. Isto é: no depósito ou arquivamento dos conteúdos. Essa ação, por sua vez, induz à adoção de uma abordagem complementar a ela, que tem sido chamada de sedução.

Sedução consiste, na verdade, em ocupar a equipe da biblioteca e outros colaboradores necessários, tais como estagiários de Biblioteconomia e áreas afins na realização de todas as atividades relacionadas com o povoamento do repositório, e, em seguida, brindar os pesquisadores com as estatísticas de uso. Isto é: trabalhar com o depósito/arquivamento mediado e, principalmente, com o *marketing* do uso. Para tanto, são necessárias ações pontuais, tais como:

- buscar, nos curriculos Lattes dos pesquisadores, na SciELO[10], no catálogo da biblioteca da universidade e em outras bases de dados ou bibliotecas digitais, dados que permitam identificar a produção científica de cada pesquisador;

10 http://scielo.org

- contatar os pesquisadores para, em conjunto com a equipe do repositório, decidir sobre os itens de sua produção a serem depositados e sobre todas as autorizações necessárias para isso;
- coletar todos os conteúdos já disponíveis em ambiente digital e sem restrições para distribuição, depositando-os imediatamente no repositório institucional;
- coletar conteúdos impressos ou quaisquer outros materiais digitais pertinentes que não foram identificados ou que possuem restrições para depósito, digitalizá-los ou reproduzi-los e em seguida depositá-los no repositório;
- divulgar, sistematicamente, estatísticas de acesso e *download* dos trabalhos depositados, como uma espécie de propaganda dos benefícios do repositório para os autores, com a expectativa de que, conhecendo os benefícios, pesquisadores se sentirão motivados a manter seus trabalhos depositados e, quiçá, passarão a realizar o autodepósito propriamente dito.

Sem dúvida nenhuma, as abordagens, estratégias e ações necessárias à criação de repositórios institucionais em universidades e institutos de pesquisa requerem criatividade e adaptação a cada contexto, como soe acontecer em relação a toda ação relacionada ao sistema de comunicação científica. No entanto, experiências bem sucedidas têm mostrado que dois pontos são fundamentais: grande empenho e conhecimento profundo das questões envolvidas com o tema acesso aberto, por parte dos *stakeholders* das iniciativas. Políticas de acesso aberto, políticas de funcionamento de repositórios, política de depósito compulsório, modelo *open archives*, direitos autorais, modelos de negócios e sustentabilidade, preservação digital, hábitos de comunicação e padrões de comportamento informacional da comunidade, são algumas das questões importantes que não se pode negligenciar.

CONTRIBUIÇÕES DOS REPOSITÓRIOS PARA AS BIBLIOTECAS DE PESQUISA

A plena adoção de repositórios institucionais em universidades e institutos de pesquisa traz consigo benefícios em todos os níveis. Benefícios para a ciência como um todo, pois, ao mesmo tempo em que cria fluxos alternativos de informação útil para a ciência, reforça e potencializa as funções da comunicação científica, o que, por sua vez, reflete positivamente no avanço e desenvolvimento científico. Benefícios para as instituições que o adotam, pois ao mesmo tempo em que instituem ou sistematizam processos coerentes de gestão da informação científica, as conectam com uma rede global e aberta de informação científica, melhorando, assim, seus processos de comunicação científica internos e externos e também a sua visibilidade. Benefícios para o pesquisador, pois, além de gerenciar e dar visibilidade à sua produção, aumentando assim o impacto dos resultados da pesquisa que realiza, contribui diretamente para o aumento da sua própria visibilidade e prestígio. Contudo, esta seção discutirá os benefícios que repositórios institucionais oferecem especificamente às bibliotecas de pesquisa, sejam elas bibliotecas de instituições de ensino e pesquisa (bibliotecas universitárias) e bibliotecas especializadas de instituições cuja finalidade principal é a produção de conhecimento para o avanço científico e tecnológico por meio de atividades de pesquisa e desenvolvimento.

Uma das principais contribuições de repositórios institucionais para as bibliotecas de pesquisa é, certamente, a melhoria do seu *status* no seio da comunidade acadêmica. Isso porque, ao se envolverem com a criação de repositórios institucionais, novos e relevantes papéis passam a ser desempenhados por seus bibliotecários, além de ser aumentada significativamente a parceria destes com pesquisadores no exercício de suas funções.

Sobre os novos papéis desempenhados, destaca-se, com base nos estudos de Lagoze e Van de Sompel (2001), o papel de editores científicos. Na verdade, bibliotecas de pesquisa, universitárias e de institutos de pesquisa, constituem, nos dias atuais, um dos principais atores dos

novos modelos de representação do processo de comunicação científica, em razão desses novos papéis, especialmente no que concerne à publicação científica. Lagoze e Van de Sompel (2001) definem como "guarda-chuva técnico da interoperabilidade" (Figura 3), a abordagem da multiplicidade de atores no processo de comunicação (COSTA; DEUSDARÁ, 2005). Nessa abordagem, bibliotecários e editores podem ser colocados lado a lado, no sentido em que realizam tarefas semelhantes de divulgação e distribuição dos resultados do trabalho de pesquisadores. Funcionam, desse modo, como mediadores do processo de comunicação entre emissores e receptores do conhecimento resultado de pesquisas.

Sobre este aspecto, sabe-se que ao longo dos tempos, as bibliotecas de pesquisa, como um dos elementos constituintes do sistema de comunicação científica, concentravam seus esforços nos processos de reunião, organização, preservação e recuperação da informação científica necessária às atividades de ensino e pesquisa desenvolvidas pelas instituições que as mantinham. Por até então ter sido considerada uma atribuição de editores científicos, bibliotecas e bibliotecários dedicavam poucos esforços a uma necessidade particular de pesquisadores: disseminar amplamente os resultados de pesquisa publicados (ou aceitos para publicação) de modo que o seu impacto fosse acelarado e aumentado. Portanto, mais do que nunca, é possível às bibliotecas e aos bibliotecários atuarem também diretamente na ampla disseminação da informação científica produzida por membros de sua instituição. O que antes era feito, quando muito, em nível institucional, por meio, por exemplo, de um sistema de disseminação seletiva da informação, é facilmente feito em escala global.

No modelo de comunicação científica reformulado pela filosofia de acesso aberto, a disseminação e distribuição da informação científica publicada (ou aceita para publicação) é uma prerrogativa compartilhada entre editores científicos e bibliotecas de pesquisa. Estas, sem dúvida alguma, tiveram seus papéis reforçados tanto na intituição quanto no sistema de comunicação científica.

Desse modo, o aumento da parceria entre bibliotecários de pesquisa e pesquisadores constitui o aspecto mais relevante relacionado à me-

lhoria do *status* de bibliotecas e bibliotecários no sistema de educação superior e de pesquisa. De fato, ao utilizarem as abordagens *top-down* (sensibilização), *bottom-up* (ação real) e arquivamento mediado com *marketing* de uso (sedução), estão, de fato, realizando uma parceria nas atividades de disseminação da pesquisa. É importante lembrar que a disseminação da pesquisa constitui o ponto alto do processo de comunicação científica que, tradicionalmente, tinha como principais atores os pesquisadores e os editores, apenas. Bibliotecas e bibliotecários de pesquisa, portanto, adquirem *status* semelhante, tornando-se, portanto, parceiros próximos dos dois. É importante mencionar também, que, ao gerir as atividades com repositórios, bibliotecários auxiliam enormemente pesquisadores nas negociações sobre autorização para depósito, junto aos editores.

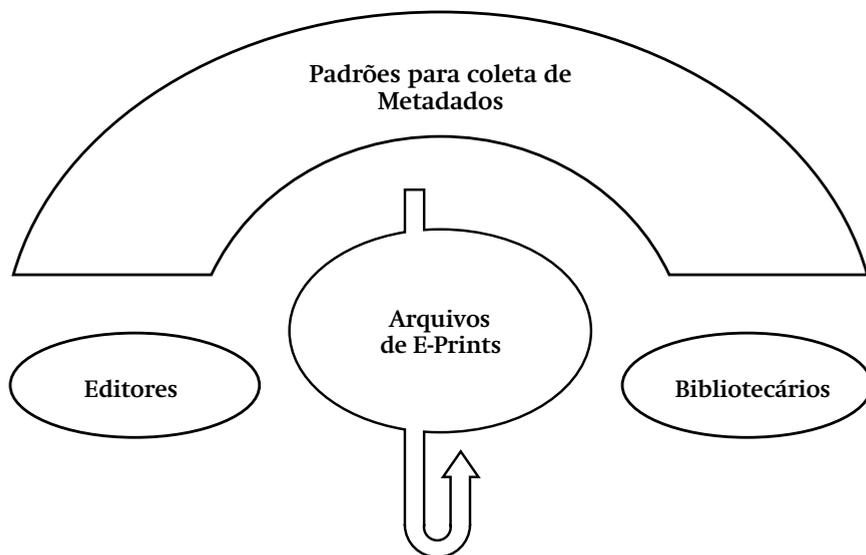



Figura 3 - Guarda-chuva técnico da interoperabilidade (multiplicidade de atores).
Fonte: Lagoze e Van de Sompel (2001).

O adequado funcionamento e a adoção efetiva de repositórios institucionais também ampliam a comunidade usuária da biblioteca de pesquisa. Ou seja, a audiência passa a ser global. Isso certamente contribui para que a importância das bibliotecas de pesquisa seja reafirmada e

evidenciada. Como já mencionado, este é um aspecto que deve ser considerado no planejamento e construção de repositórios institucionais.

Nas recomendações descritas nos resultados obtidos com os trabalhos da Budapest Open Access Initiative (BOAI, 2002) e muito bem ressaltadas por Suber (2003), bibliotecários devem "oferecer-se para manter repositórios em suas instituições, ajudando os docentes-pesquisadores a depositar seus trabalhos anteriores, digitalizá-los, se necessário, e ensiná-los a depositar trabalhos futuros". De fato, claramente a recomendação de depósito mediado.

O *marketing* realizado com a divulgação das estatísticas de uso dos itens depositados é uma espécie de testemunho dos benefícios dos repositórios, benefícios esses alcançados por meio do trabalho de bibliotecários em bibliotecas de pesquisa. As observações de Crow (2002a) sobre o assunto mostram que o trabalho com repositórios institucionais introduz mudanças significativas no sistema de comunicação científica, tais como:

- expandem exponencialmente o acesso à pesquisa;
- reafirmam o controle sobre o saber pela academia;
- reduzem o monopólio dos periódicos científicos (especialmente no que concerne à acessibilidade apenas por assinatura);
- servem como indicadores tangíveis da qualidade de uma universidade;
- demonstram a relevância científica, social e econômica das atividades de pesquisa da instituição;
- aumentam a visibilidade, o *status* e o valor público da instituição.

Fundamentados nessas questões, Leite e Costa (2006) propõem uma abordagem segundo a qual repositórios institucionais constituem ferramentas apropriadas de gestão do conhecimento científico no contexto das universidades. Nessa perspectiva, Lynch (2003) define repositórios institucionais de universidades como "um conjunto de serviços que a universidade oferece aos membros de sua comunidade, visando ao gerenciamento e disseminação dos materiais digitais criados pela instituição e pelos membros de sua comunidade". Repositórios institucionais podem, portanto, ser vistos como ferramentas

adequadas para a gestão do conhecimento científico, pois, ao mesmo tempo em que agilizam os processos de comunicação científica, potencializam a condução de processos de gestão, maximizando a criação, o compartilhamento, a disseminação e o uso do conhecimento científico.

Resultados de um estudo realizado por Antelman (2005) mostram que bibliotecários devem trabalhar nesse tipo de iniciativas. Para isso, devem criar repositórios institucionais, buscar alternativas para publicações com acesso aberto e auxiliar docentes pesquisadores nas universidades a negociar direitos autorais com editores.

Outro aspecto relevante é uma nova função institucional assumida pelas bibliotecas de pesquisa. A gestão da informação científica institucional, especialmente aquela publicada, é uma atribuição recentemente assumida pelas bibliotecas de pesquisa, principalmente pelas universitárias. Até então, bibliotecas universitárias gerenciavam acervos quer serviam de insumo para as atividades de ensino e pesquisa. Em relação à produção científica institucional os esforços alcançavam, quando muito, as teses e dissertações da instituição, diferentemente das bibliotecas de institutos de pesquisa, que, em geral, já possuíam controle referencial de sua produção científica institucional. Por motivos próprios do modelo tradicional de comunicação científica que promove a falta de controle da instituição sobre sua produção científica, bibliotecas de pesquisa até então não eram capazes de oferecer acesso sistematizado e integral à informação científica produzida por membros de sua instituição. Repositórios institucionais possibilitaram que essas bibliotecas fossem além disso, ou seja, mais do que acesso, ofereçam indicadores que contribuem para avaliação das atividades de pesquisa e desenvolvimento das instituições.

É claro que há muito mais. Estas são apenas algumas reflexões importantes que visam a chamar a atenção para as mudanças significativas introduzidas no sistema de comunicação científica nos dias atuais, por meio do trabalho de bibliotecários. Consequentemente, não deve restar nenhum equívoco sobre os novos e importantes papeis por eles desempenhados, nem sobre o status ricamente elevado das bibliotecas em que atuam.

UM POUCO SOBRE A EXPERIÊNCIA INTERNACIONAL COM REPOSITÓRIOS INSTITUCIONAIS

Como explicitado no início deste capítulo, a história dos repositórios institucionais é recente, e se iniciou, como soe acontecer à grande maioria das inovações na área de informação, no hemisfério Norte. Expandiu-se, no entanto, rapidamente, como também soe acontecer a inovações tecnológicas, embora, igualmente, no hemisfério Norte. Os Estados Unidos lideram mais uma vez as iniciativas, embora a Europa tenha um número significativo de exemplos a serem observados, ou mesmo seguidos. A figura 4 ilustra a distribuição de repositórios digitais, dentre eles os institucionais, no globo.

O Directory of Open Access Repositories – *Open*DOAR[11] é o principal diretório de registro de repositório digitais de acesso aberto. Atualmente, o diretório contabiliza 1.583 repositórios digitais registrados. Desse total, 1.278 correspondem a repositórios institucionais propriamente ditos. É importante considerar que a quantidade de repositórios presentes no *Open*DOAR não expressa fidedignamente a realidade, pois, por se tratar de um registro voluntário, naturalmente, o número de repositórios é maior.

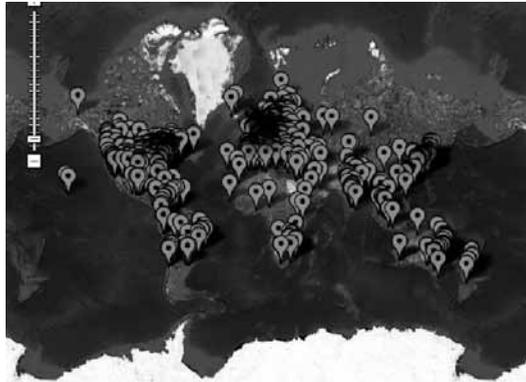



Figura 4 - Repositórios digitais distribuídos no mundo
Fonte: *Open*DOAR

11 http://www.opendoar.org/

O crescimento do *Open*DOAR (Figura 5) constitui uma clara evidência da emergência e consolidação do acesso aberto como modelo alternativo de comunicação científica, que tem na Via Verde uma poderosa estratégia. O diretório registra um crescimento constante da quantidade de repositórios digitais registrados ao longo dos seus quatro anos de existência, uma média de 1,08 repositórios registrados por dia.

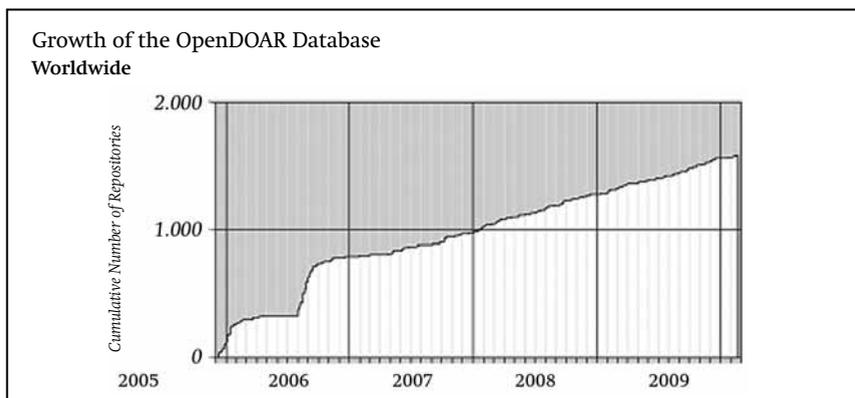



Figura 5: Crescimento do *Open*DOAR.
Fonte: OpenDOAR.

Embora a América do Norte e a Europa ainda concentrem maior parte dos repositórios digitais, conforme dados extraídos do *Open*DOAR, a Via Verde já pode ser considerada uma realidade internacional (Figuras 6 e 7).

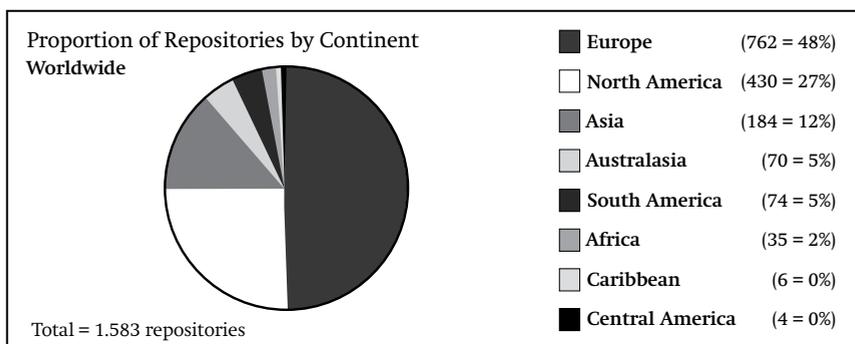



Figura 6: Proporção de repositórios por continente.
Fonte: *Open*DOAR

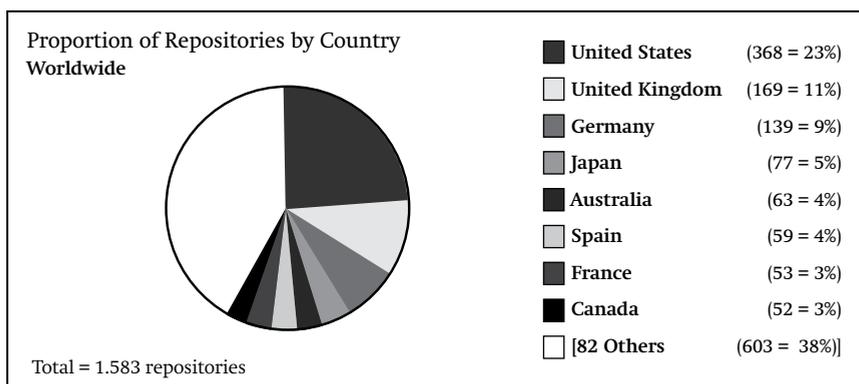



Figura 7: Proporção de repositórios por país.
Fonte: *Open*DOAR.

Rieh e colaboradores (2007) realizaram censo de repositórios institucionais no Estados Unidos, país de onde na realidade emanam os principais desenvolvimentos do acesso aberto. O levantamento revelou que a maioria das instituições que implementaram repositórios institucionais foi de universidades. Ao concluir que repositórios institucionais servem a diversas propostas nas diferentes categorias de instituições relacionadas com o ensino superior, o autor recorre à Lynch e Lippincott (2005), que identificaram dois tipos de repositórios institucionais, conforme estudo realizado nos Estados Unidos. O primeiro tipo é àquele que assume a função de instrumento para a disseminação de *e-prints* produzidos por membros da instituição. O segundo tipo, por sua vez, diz respeito aquele que constitui acervo do trabalho intelectual da instituição resultante tanto das atividades de pesquisa quanto de ensino. A tendência identificada nos Estados Unidos é refletida no conjunto de dados extraídos do *Open*DOAR e ilustrados na figura 8. Com base em informações fornecidas no momento do registro do repositório no diretório, o gráfico explicita os tipos de conteúdos que constituem as coleções dos repositórios registrados, boa parte relacionada com atividades de ensino e pesquisa.

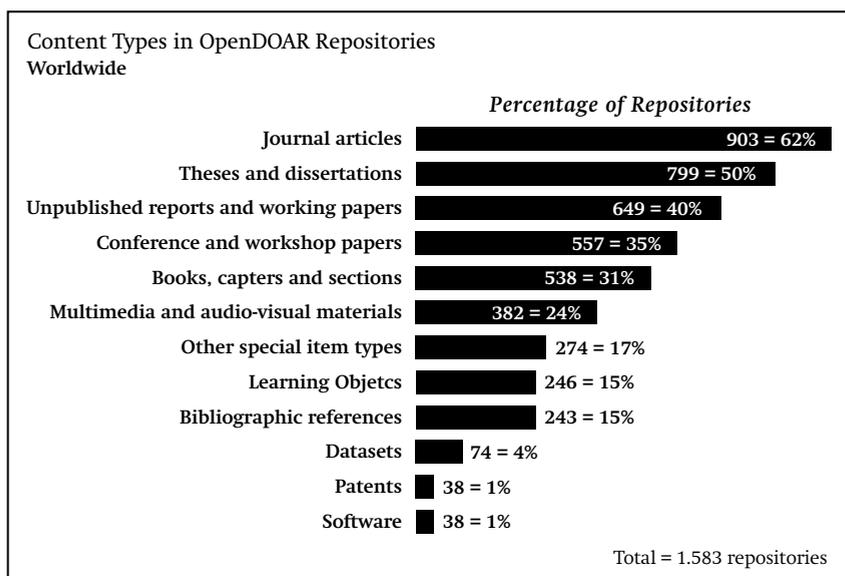



Figura 8 - Tipos de conteúdos de repositórios do *Open*DOAR
Fonte: *Open*DOAR.

A despeito da ênfase que parte da literatura especializada no tópico costuma dar ao depósito da produção científica avaliada por pares (o que significa a preponderância de artigos de periódicos científicos), os dados revelam diversidade de tipos de conteúdos. É possível que tal fenômeno ocorra também por influência de demandas institucionais que possuem relação apenas indireta com as funções da comunicação científica. Quer se dizer com isso que, dado o potencial dos repositórios institucionais como ferramentas adequadas para a gestão da informação em universidades e instituições de pesquisa, conteúdos de natureza acadêmica que não contêm, necessariamente, resultados de pesquisa, passam a constituir, por necessidade institucional, objeto de trabalho das bibliotecas que gerenciam os repositórios. Isso porque esses conteúdos são também produzidos intensivamente por seus membros – como material de ensino e aprendizagem – e que até então careciam de processos de gestão apropriados. Ou seja, repositórios institucionais tornaram-se uma oportunidade para a integração de processos e de sistemas relacionados tanto com a pesquisa quanto com o ensino.

Na realidade, em todo o mundo percebe-se uma forte associação entre a pesquisa e o ensino. Reforçando este argumento, Godin e Gringras (2000) com base em resultado de pesquisa realizada no Canadá, afirmam que a despeito de uma diversificação real dos loci de produção, universidades ainda estão no coração do sistema e todos os outros atores contam pesadamente com a *expertise* delas. Entenda-se que outros atores incluem institutos de pesquisa privados ou públicos, indústrias, laboratórios, hospitais e outras instituições participantes do sistema de produção do conhecimento. Nesse sentido, os autores apresentam uma série de dados que indicam o aumento da quantidade de trabalhos científicos, em relação ao total de trabalhos científicos publicados no Canadá, cuja autoria provém de universidades. Desse modo, desde que conduzida da maneira adequada, seja em um único ou em mais de um repositório na instituição, a flexibilização dos critérios de constituição de suas coleções digitais é oportuna e enriquecedora das atividades de produção do conhecimento por meio da integração de recursos de informação resultantes do ensino e da pesquisa.

SITUAÇÃO DOS REPOSITÓRIOS INSTITUCIONAIS DA UNIÃO EUROPEIA

O mais recente e detalhado relato da situação dos repositórios institucionais na União Europeia foi feito por Van der Graaf e Van Eijndhoven (2008), em estudo financiado pelo *DRIVER.* Com objetivo de evidenciar um panorama geral de aspectos relacionados com desenvolvimento e funcionamento de repositórios institucionais na Europa, destaca-se a seguir alguns dos resultados do levantamento.

- o número de repositórios institucionais de acesso aberto à informação científica é estimado entre 280 e 290, o que significa que quase metade das universidades possui agora repositórios institucionais implementados. O crescimento girou em torno de 25 e 30 repositórios por ano ao longo de três anos;
- em relação a tipos de conteúdos, 90.4% dos repositórios analisados contêm artigos de periódicos, 80.3% contêm teses e dissertações, 69.7% contêm *working papers*, 74.7% contêm livros ou capítulos de

livros, 67.4% contêm anais de conferências. Sobre materiais não textuais, os autores relatam que 21.3% contêm imagens, 18.0% contêm vídeos, 8.4% dados de pesquisa, 3.9% contêm música e 33.1% relataram outros tipos de conteúdos como patentes e materiais de aprendizagem. A minoria dos repositórios contém apenas metadados dos documentos sem os textos completos;

- no que diz respeito à versão dos artigos científicos depositados nos repositórios institucionais, 71.9% permitem a versão publicada pelas editoras, 70.8% a versão pós-print e 50% a versão pré-print. Sobre a quantidade de cada uma das versões nos repositórios, 43.6% dos respondentes afirmaram que possuem principalmente a versão publicada, 46.2% o pós-print e 10.3% o pré-print;
- em relação ao acesso aberto aos conteúdos e outras modalidades de disponibilidade, 96.6% afirmaram conter materiais textuais disponíveis sem restrições de acesso, 32.6% acesso aberto submetido a período de embargo, 30.0% com acesso restrito à instituição, 18.0%, depositados mas sem a liberação do acesso e 6.7% acessível de uma outra forma como, por exemplo, o pagamento de taxas e restrição à determinados grupos. Sobre as quantidades de cada uma das modalidades presentes no repositório, 86% dos conteúdos acessíveis integralmente, 3% dos conteúdos submetidos a embargo, 5% acessíveis somente na instituição, 4% dos conteúdos depositadados mas não acessíveis de nenhuma maneira e 1% acessível sob alguma outra condição;
- sobre a modalidade de depósito praticado nos repositórios institucionais analisados, i) 21.3% afirmaram que praticam o autoarquivamento pelos próprios autores com controle de qualidade por pessoal especializado, ii) 20.2% afirmaram que autores entregam o texto completo e a equipe do repositório efetua o depósito, iii) 8.4% coletam os conteúdos independemente dos autores, 43.8% admitem a combinação de i, ii e iii, 6.2% informaram outros tipos de procedimentos (autoarquivamento sem controle de qualidade pela biblioteca e outras combinações);
- sobre os *software* utilizados, 30.3% adotaram o Dspace, 19.7% o

Eprints. Os 50% restantes adotaram outras ferramentas: 5% ARNO, CDSWare, Digitool, DIVA, Fedora, iTOR, MyCoRe e VITAL), 16.9% ferramentas desenvolvidas localmente, 11.2% outros pacotes;

- 84.3% assinaram identificadores persistentes para cada item do repositório, 12% não adotaram qualquer indentificador persistente e 4% afirmaram desconhecer;
- 71.9% dos repositórios registram dados estatísticos de acesso e uso dos conteúdos, 21.3% não registram e 6.7% afirmaram desconhecer o assunto;
- em relação aos padrões de metadados adotados, 61.8% adotam o Dublin Core e qualificadores, 43.3% apenas o Dublin Core, 9% o DIDL, 11.2% o MARC21, 12.4% o MODS 14.6% o OAI-ORE;
- sobre as políticas relacionadas com o depósito de documentos pelos autores, 29.2% afirmaram ter uma política de depósito voluntário combinada com forte encorajamento para adotá-lo, 23% adotam apenas uma política de depósito vonluntário, 20.2% possui uma política de depósito parcialmente obrigatório (os autores são obrigados a depositar suas teses, para outros tipos de conteúdos o depósito é opcional), 11.8% possuem uma política de depósito obrigatório e 14.6% não adotaram nenhuma política formal relacionada com o depósito de conteúdos;
- questionados a respeito de serviços adicionais, um dos aspectos interessantes das repostas versa sobre a presença do repositório em catálagos. Dos respondentes, 47.8% têm seus conteúdos listados nos catálogos da biblioteca da instituição, 44.4% têm seus conteúdos listados em catálogos regionais ou nacionais e 11.8% em outros catálogos.
- em relação aos fatores de maior estímulo para o desenvolvimento de repositórios institucionais e seus conteúdos, 30% dos respondentes apontaram o aumento da visibilidade e citações, interesse dos tomadores de decisão da instituição, o simples processo de depósito do ponto de vista dos depositantes e a integração do repositório com outros sistemas na instituição. Entre 20% e 30% indicaram que a exigência das agências de fomento em seus países para o depósito dos resultados de pesquisa financiada em

repositórios de acesso aberto, esforços de sensibilização de pesquisadores além da existência de políticas de depósito obrigatório. Por fim, entre 10% e 20% apontaram a política de salvaguarda e preservação digital de longo prazo do acervo do repositório e também a situação dos direitos de cópia de materiais publicados e o conhecimento dos autores sobre esta questão;

- por outro lado, 50% dos respondentes apontaram como principais fatores inibidores a falta de uma política institucional de depósito obrigatório, a situação de direitos de cópia de materiais publicados e o conhecimento dos autores sobre esta questão. Entre 20% e 30% apontaram a falta de exigência das agências de fomento em seus países para a obrigatoriedade dos resultados de pesquisa em repositórios de acesso aberto, a falta de integração do repositório com outros sistemas na instituição e o desinteresse dos tomadores de decisão da instituição. Entre 10% e 20% indicaram a falta de esforços de sensibilização de pesquisadores em sua instituição e a carência de suporte para o aumento da visibilidade e citações da produção científica dos pesquisadores. Menos de 10% apontaram a falta de uma política de preservação digital de longo prazo como um fator inibidor.

Essa espécie de fotografia da experiência internacional, embora restrita aos Estados Unidos e à Europa, dá o tom da *praxis* que envolve a questão dos repositórios. Isso não significa que outras regiões tais como Ásia e América Latina não estejam, do mesmo modo, inseridas nesse contexto, como já registrado na literatura.

É bom lembrar que esta é uma prática de menos de uma década, mas que já causou impactos em toda a comunidade científica, em dois sentidos. Primeiro a totalidade da comunidade científica em relação à sua distribuição no globo. Segundo, a totalidade da comunidade científica em relação aos tipos de atores dela constituintes. Repositórios institucionais tornam-se, desse modo, elo fundamental nas novas cadeias do sistema de comunicação científica.

USO DE REPOSITÓRIOS INSTITUCIONAIS E A VISIBILIDADE DOS RESULTADOS DE PESQUISA: ALGUNS EXEMPLOS SIGNIFICATIVOS

Inúmeros estudos têm sido realizados para mensurar a relação entre o aumento das taxas de citação da literatura científica e sua disponibilidade em ambiente de acesso aberto (LAWRENCE, 2001; KURTZ, 2004; HARNAD, BRODY, 2004a; ANTELMAN, 2004; NORRIS et al, 2008; BRODY, 2004; KOUSHA, ABDOLI, 2009; DAVIS et al 2008). De um modo geral, a mensuração de impacto de resultados de pesquisa publicados, como ressalta Swan (2008), é realizada a partir de métricas que podem ser de dois tipos: i) métricas geradas pelo autor, análises baseadas nas citações (citações formais, agradecimentos, *links* e outros) e ii) métricas geradas pelos usuários (leitores), cujas análises são feitas tendo por base o uso (*downloads*, visualização de metadados e outros).

A correlação entre citações e *downloads* é explorada por Brody e colaboradores (2006). Por conta do aumento de artigos científicos acessados por meio da Web, é possível registrar e contabilizar instantaneamente o número de vezes que foi descarregado (*downloads*). Ainda segundo os autores, é de se esperar que o número de vezes que um artigo foi lido esteja relacionado tanto com o seu tempo de publicação quanto com a quantidade de vezes que ele foi citado.

Repositórios institucionais podem gerar indicadores de uso e mostrar a quantidade de *downloads* dos conteúdos que armazenam, conforme apontam Swan e Chan (2009). Esta seção apresentará alguns dados estatísticos de uso de repositórios institucionais que evidenciam os benefícios para a disseminação da informação científica.

Um caso mencionado na literatura é o de um pesquisador australiano da área de Química da Queensland University of Technology, Ray Frost, que possui atualmente 537 documentos de sua autoria no QUT ePrints[12]. No período de um ano (entre 20/02/2009 e 20/02/2010) todos os documentos do pesquisador receberam um total de 189.856 de *downloads*, o que lhe torna o autor cujos textos foram mais acessados no repositório institucional (Tabela 1).

12 http://eprints.qut.edu.au/

Tabela 1 - Os 10 autores cujos textos foram mais acessados no QUT ePrints

Rank	Autor	Total
1	Frost, Ray	189856
2	Skitmore, Martin	187229
3	Ferreira, Luis	125936
4	Worthington, Andrew	124015
5	Watters, James	117348
6	Courtney, Mary	117111
7	Fitzgerald, Brian	113445
8	Cunningham, Stuart	112100
9	Campbell, Marilyn	105655
10	Diezmann, Carmel	104757

Fonte: QUT ePrints.

Seu artigo intitulado *Porous materials for oil spill cleanup: a review of synthesis and absorbing properties*, publicado pelo *Journal of Porous Materials*, um periódico científico da editora Springer, em 2003, foi depositado no repositório em junho de 2005. Somente no período em destaque, o texto recebeu 6.286 *downloads*. O QUT ePrints ocupa o 19° lugar no *ranking* mundial de repositórios institucionais, que mensura a visibilidade global e o impacto dos repositórios.

Outro depoimento significativo é o do Repositório Institucional da Universidad de los Andes[13], na Venezuela. No período compreendido entre setembro de 2008 e julho de 2009 o repositório registrou 2.174.697 *downloads* (Figuras 9 e 10). Nesse período o repositório contava com 19.983 documentos, alcançando uma média de 108,8 *downloads* por item.

13 http://www.saber.ula.ve/

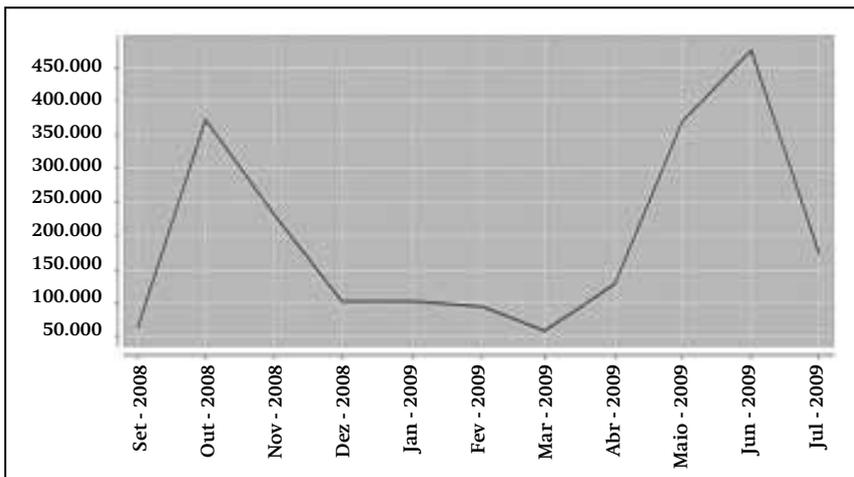



Figura 9 - Quantidade de *downloads* por mês e ano.
Fonte: SABER - ULA

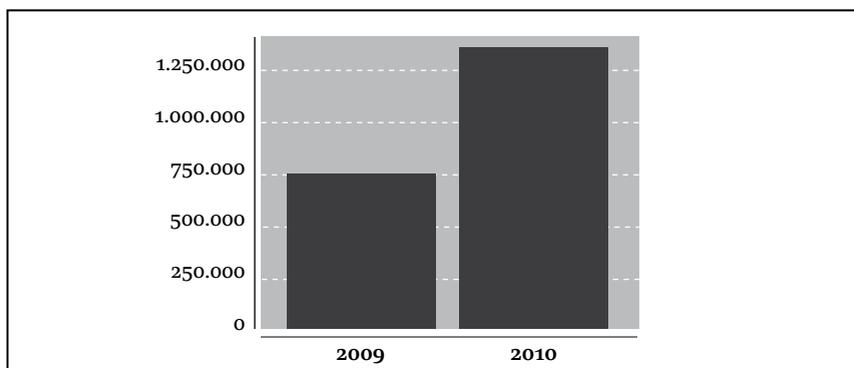



Figura 10 - Quantidade de *downloads* por ano.
Fonte: SABER - ULA

O RepositoriUM[14], repositório institucional da Universidade do Minho, é considerado um exemplo de iniciativa cujos benefícios são claramente percebidos. Desde janeiro de 2006 até fevereiro de 2010 o RepositoriUM registrou 3.772.545 de *downloads*, e seu acervo conta com 9.262 itens, uma média de 407,3 *downloads* por item (Figura 11). A figura 12 mostra a quantidade de *downloads* por país. A tabela 2 apresenta dados estatísticos de uso de diversos repositórios institucionais.

14 http://repositorium.sdum.uminho.pt

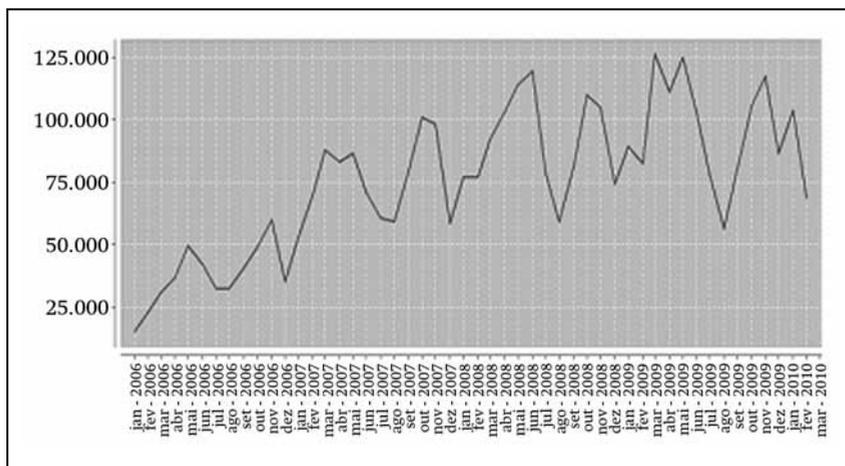



Figura 11 - Quantidade de *downloads* por mês e ano.
Fonte: RepositoriUM.

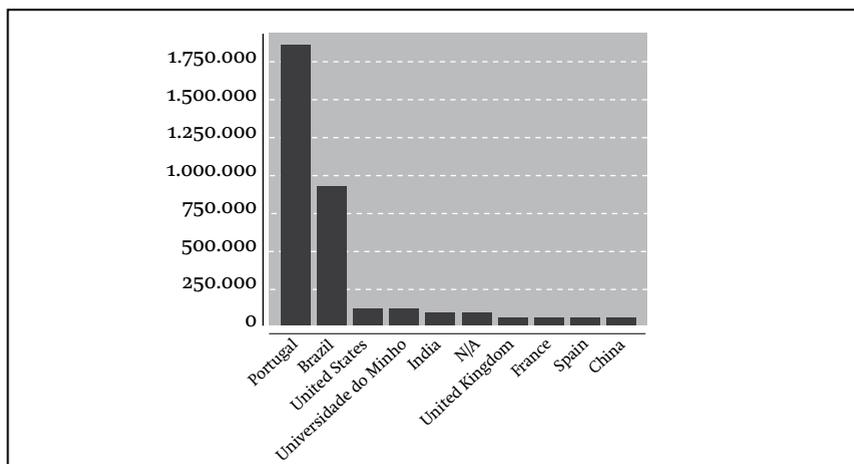



Figura 12 - Quanttidade de *downloads* por país.
Fonte: RepositoriUM.

Tabela 2 - Estatísticas de uso de repositórios institucionais

Repositório Institucional	Quantidade de itens	Número de downloads	Downloads em 2009
UNL's Institutional Repository	38,922	3,556,992	1,524,287
eCommons@Cornell	12,478	7,418,581	780,036
Purdue e-Pubs	12,931	740,616	307,452
University of Calgary Institutional Repository	17,202	1,369,282	676,823
ScholarlyCommons Repository at the University of Pennsylvania	12,816	1,472,057	441,057
ScholarSpace at University of Hawaii at Manoa	10,522	299,031	223,925
Research Online at the University of Wollongong	7,437	1,248,993	516,560
DigitalCommons@UTEP	9,358	349,702	46,838

Fonte: repositórios institucionais citados.

Em apresentação realizada na *American Library Association Annual Convention*, Royster (2009) comenta exemplos da visibilidade de conteúdos depositados no University of Nebraska – Lincoln Institutional Repository[15], durante o mês de março de 2009, reproduzidos a seguir:

- 77% de todo o acervo do repositório institucional (17.460 itens) registraram algum (*download*) uso nesse período, e apenas 23% (5.320 itens) não registraram nenhuma estatística de uso;
- foram registrados 137.072 *downloads*, sendo que nesse período o repositório contava com cerca de 21.000 itens;
- nesse mesmo período o Departamento de Psicologia contava com 367 itens depositados no repositório e o registro de 5.008 *downloads*, uma média de 13.6 *downloads* por item;

15 http://digitalcommons.unl.edu/

- as 190 publicações de autoria de Robert Katz, pesquisador aposentado em 1987, registraram, nesse mesmo período, 1.357 *downloads*, uma média de 7.1 por item;
- o repositório tinha depositado 2.274 documentos dos *Larsen Tractor Museum Archives*, acumulando, até então 16.648 *downloads* em março de 2009, uma média de 7.3 *downloads*;
- nesse mesmo período havia no repositório 6 dissertações nas coleções do Departamento de Língua e Literatura Moderna, 984 *downloads* e uma média de 164 por item;
- o *Online Dictionary of Invertebrate Zoology*, de Mary Ann Maggenti e colaboradores, publicado em 2005, foi o documento mais descarregado nesse período. O dicionário recebeu 824 *downloads* em sua versão completa, e o conjunto de arquivos separados (24 capítulos) recebeu 1.903 *downloads*;
- 26% do total de *downloads* recebidos nesse período (36.000) foram originados de outros países (147 países).

Portanto, é fundamental que o planejamento de repositório institucionais preveja o uso de pacotes de estatísticas. Além de oferecer indicadores para a gestão das atividades de pesquisa e desenvolvimento da instituição, dados estatísticos gerados por repositórios institucionais são imprescindíveis para a demonstração de benefícios do acesso aberto nas estratégias de convencimento de pesquisadores para o depósito da produção científica.

REFERÊNCIAS

ANTELMAN, K. Do open access articles have a greater research impact? **College & Research Libraries**, v. 65, n. 5, p. 372-382, maio 2004.

BARTON, M. R. Digital institutional repository development: issues and solutions. In: COSTA, S. M. S. et al. **From information to knowledge**: Proceedings of the 7th ICCC/IFIP International Conference on Electronic Publishing, Universidade do Minho, 2003. Disponível em: <http://elpub.scix.net/cgi-bin/works/Show?0303.>.

BARTON, M. R.; WATERS, M. M. **Creating an institutional repository**. [Cambridge]: MIT, 2004. 134 p. LEADIRS workbook. Disponível em: <http://

www.dspace.org/implement/leadirs.pdf>.

BOAI - Budapest Open Access Initiative. 2002. Disponível em: <http://www.soros.org/openaccess>.

BRODY, T. Citation analysis in the open access world. **Interactive Media International**, 2004. Disponível em: <http://eprints.ecs.soton.ac.uk/10000/1/tim_oa.pdf>.

BRODY, T. et al. Earlier Web usage statistics as predictors of later citation impact. **Journal of the American Association for Information Science and Technology**, v. 57, n. 8, 2006, p. 1060-1072. Disponível em: <http://eprints.ecs.soton.ac.uk/10713/>

COSTA, S. M. S.; DEUSDARÁ, A. P. Brazilian open archives initiatives: the influence of ElPub conferences and the role of Ibict. In: DOBREVA, M.; ENGELEN, J. (Ed.). **From author to reader** challenges of the digital content chain: proceedings of the 2005 Elpub conference. Leuven: Peeters Publishers, 2005.

CROW, R. **The case for institutional repositories**: a SPARC position paper. Washington, DC: ARL, 2002a. Disponível em: <http://www.arl.org/sparc/bm~doc/ir_final_release_102.pdf>

CROW, R. **Institutional repository**: checklist and resource guide. Washington, DC: SPARC, 2002b. 51 p. Disponível em: <www.arl.org/sparc/bm~doc/IR_Guide_&_Checklist_v1.pdf>

DAVIS, P. M. et al. Open access publishing, article downloads, and citations: randomised controlled trial. **British Medical Journal**, 337, a568, p. 1-6, 2008. Disponível em: <http://www.bmj.com/cgi/doi/10.1136/bmj.a568>

GODIN, B.; GINGRAS, Y. The place of universities in the system of knowledge production. **Research Policy**, v. 29, 2000, p. 273-278. Disponível em: <http://econpapers.repec.org/article/eeerespol/v_3a29_3ay_3a2000_3ai_3a2_3ap_3a273-278.htm>

HARNAD, S.; BRODY, T.Comparing the impact of open access (OA) vs. non-OA articles in the same journals. **D-Lib Magazine**, v. 10, n. 6, June 2004. Disponível em: <http://www.dlib.org/dlib/june04/harnad/06harnad.html>

HARNAD, S. et al. **The access/impact problem and the green and gold roads to open access**. 2001. Disponível em: <http://www.ecs.soton.ac.uk/~harnad/Temp/impact.html>

JONES. R. et al. **The institutional repository**. Oxford: Chandos Publishing, 2006. 268 p.

KOUSHA, K.; ABDOLI, M. The citation impact of open access agricultural research: a comparison between OA and non-OA publications. In: WORLD LIBRARY AND INFORMATION CONGRESS: IFLA GENERAL CONFERENCE AND COUNCIL, 75., 2009, Milan, Italy. Proceedings… Milan: IFLA, 2009. Disponível em: <http://www.ifla.org/files/hq/papers/ifla75/101-kousha-en.pdf>

KURTZ, M. J. Restrictive access policies cut readership of electronic research journal articles by a factor of two. Harvard-Smithsonian Centre for Astrophysics, Cambridge, MA. **National Policies on Open Access (OA) Provision for University Research Output**: an International meeting, Southampton, 19 February 2004. Disponível em: <http://opcit.eprints.org/feb190a/kurtz.pdf>

LAGOZE, Carl; VAN DE SOMPEL, Herbert. **The open archives initiative**: building a low-barrier interoperability framework. 2001. Disponível em: <http://www.openarchives.org/documents/jcdl2001-oai.pdf>

LEITE, F. C. L. **Como gerenciar e ampliar a visibilidade da informação científica brasileira**: repositórios institucionais de acesso aberto. Brasília: IBICT, 2009. 120p. Disponível em: <http://eprints.rclis.org/17190/>

LEITE, F. C. L.; COSTA, S. M. S. Repositórios institucionais como ferramentas de gestão do conhecimento científico no ambiente acadêmico. **Perspectivas em Ciência da Informação**, v. 11, n. 2, 2006. Disponível em:< http://www.eci.ufmg.br/pcionline>

LYNCH, C. A. Institutional repositories: essential infrastructure for scholarship in the digital age. **ARL Bimonthly Report**, v. 26, 2003. Disponível em: <http://www.arl.org/newsltr/226/ir.html>

LYNCH, C. A.; LIPPINCOTT, J. K. Institutional repository deployment in the United States as of early 2005. **D-Lib Magazine**, v. 11, n. 9, Sep. 2005. Disponível em: < http://www.dlib.org/dlib/september05/lynch/09lynch.html>.

MCKAY, D. Institutional repositories and their 'other' users: usability beyond authors.
Ariadne, n. 52, 2007. Disponível em: <http://www.ariadne.ac.uk/issue52/mckay/>

NORRIS, M., et al. The citation advantage of open access articles. **Journal of the American Society for Information Science and Technology**, v. 59, n. 12, 2008, p.1963–1972. Disponível em: <http://www3.interscience.wiley.com/cgi-bin/fulltext/120748494/PDFSTART>

PROUDMAN, V. The population of repositories. In: WEENINK, K., et al. **A DRIVER's guide to European repositories**. Amsterdam: Amsterdam University Press, 2008. p. 49-102. Disponível em: <http://dare.uva.nl/document/93898>

ROOSENDAAL, H. E.; GEURTS, P. A. T. M. **Forces and functions in scientific communication**: an analysis of their interplay. 1998. Disponível em: <http://www.physik.uni-oldenburg.de/conferences/crisp97/roosendaal.html>

SPARC. **Open doors and open minds**: what faculty authors can do to ensure open access to their work through their institution. Cambridge, 2008. Disponível em: <http://www.arl.org/sparc/bm~doc/opendoors_v1.pdf>

SWAN, A. **New metrics for research outputs**: overview of the main issues.

Technical Report UNSPECIFIED, School of Electronics & Computer Science, University of Southampton, 2008. Disponível em: <http://eprints.ecs.soton.ac.uk/16986/>

SWAN, A. The business of digital repositories. In: WEENINK, K., et al. **A DRIVER's guide to European repositories**. Amsterdam: Amsterdam University Press, 2008. p. 15-48. Disponível em: <http://dare.uva.nl/document/93898>

VAN EIJNDHOVEN, K.; VAN DER GRAAF, M. **Inventory study into the present type and level of OAI compliant Digital Repository activities in the EU**. 2008. Disponível em: <http://dare.uva.nl/document/93725>

VAN DER GRAAF, M.; VERNOOY-GERRITSEN, M. **The european repository landscape 2008: inventory of digital repositories for research output in the EU.** Amsterdam: Amsterdam University Press, 2009. 169p. Disponível em: <http://dare.uva.nl/document/150724>

WEENINK, K., et al. **A DRIVER's guide to european repositories**. Amsterdam: Amsterdam University Press, 2008. Disponível em: <http://dare.uva.nl/document/93898>

Repositórios institucionais:
políticas e mandatos

Hélio Kuramoto

INTRODUÇÃO

Os Repositórios Institucionais (RI) constituem, hoje, uma das principais iniciativas para a implantação do acesso livre no mundo. Os benefícios que essa iniciativa traz para os pesquisadores, as instituições de ensino e pesquisa, assim como para os países são expressivos.

Para os pesquisadores, essa iniciativa proporciona maior visibilidade[1] às suas pesquisas e, mesmo à sua pessoa. Além da possibilidade de ter o seu trabalho citado, ele, certamente, ficará mais conhecido no meio acadêmico. Da mesma forma, as instituições de ensino e pesquisa que mantém os seus RI atualizados ganham também maior visibilidade e competitividade. Para os países, a construção de uma rede de RI integrando a sua produção científica nacional, significa dar maior transparência aos investimentos em ciência, assim como maior governabilidade. Esses benefícios aos países ou governos é possível graças às possibilidades que estes ganham com a implantação de RI em todas as suas universidades e centros de pesquisa. Mais do que simplesmente promover o registro e a disseminação da sua produção científica nacional, a criação de uma rede nacional de RI possibilita a geração de uma série de indicadores que serão de fundamental importância na gestão de financiamento da ciência em um país.

O desenvolvimento e implantação de RI, em princípio, hoje, não apresentam grandes dificuldades técnicas ou tecnológicas, uma vez

1 Veja artigo de Alma Swan (2010, p. 167), onde a autora mostra resultados de estudos realizados por Stevan Harnad e sua equipe.

que existem pacotes de software *open source* que permitem o seu desenvolvimento e implantação e oferecem facilidades para a sua construção de forma a adequar as necessidades de informação de cada instituição. Além disso, existem na internet diversos textos contendo metodologias de desenvolvimento de um repositório e, também, documentos contendo diversas iniciativas e experiências.

Apesar das facilidades citadas, ainda assim encontramos dificuldades na implantação de um RI. Essas dificuldades dizem respeito ao seu processo de gestão. O desenvolvimento e implantação de RI dependem dos objetivos deste repositório e os objetivos definem que tipos de documentos se desejam armazenar no mesmo. E cada tipo de documento tem suas características próprias.

Com o surgimento das tecnologias da informação e da comunicação, dentre uma série de outras mudanças provocadas por elas, está a forma de entrada de documentos em um RI. Em passado não muito longínquo, a entrada de dados se fazia de forma centralizada, ou por intermédio dos digitadores, nas décadas de 1960, 1970 e 1980, ou por meio de pessoas especializadas como os bibliotecários, mais recentemente.

Desde o final do século passado, mais especificamente desde o final da década de 90, o processo de entrada de dados vem se descentralizando graças às novas tecnologias da informação e da comunicação que se popularizam cada vez mais. Quem não se lembra da entrega das declarações de imposto de renda de pessoa física em disquete? E, mais recentemente, desde o início de 2000 a maioria das declarações de imposto de renda de pessoa física são entregues via internet. Esta modalidade de entrada de dados funciona porque o cidadão é obrigado a entregar a sua declaração de imposto de renda e, eventualmente tem uma restituição a receber. Existe, portanto, um mandato ou obrigatoriedade estabelecida em lei e, além disto, existe uma penalização na forma de multa caso a declaração não seja entregue. Aliado a estes fatos, deve-se considerar também que é muito mais fácil preparar a declaração de imposto de renda em meio eletrônico do que em papel, além da agilidade e rapidez em um eventual recebimento de restituição. Esta forma de entrada de dados foi e é tão eficaz que a

Receita Federal está prevendo, para o ano de 2011, abolir a entrega de declaração de imposto de renda em papel.

Outro exemplo de sucesso relacionado com a entrada distribuída de dados é a Plataforma Lattes, no qual os pesquisadores que pleiteiam um auxílio à sua pesquisa são obrigados a criar e manter atualizados os seus respectivos currículos. A não atualização de seus currículos tem como reflexo a não concessão dos auxílios pleiteados. Portanto, neste caso, também existe uma moeda de troca ou uma política de atualização dos currículos Lattes.

Da mesma forma, os diversos sistemas de gestão de RI (Dspace, Eprints etc.) se baseiam na entrada de dados via autodepósito ou autoarmazenamento, ou seja, o próprio autor é quem faz o depósito de seu documento nesses sistemas.

É importante ressaltar que o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), ao implantar a Biblioteca Digital Brasileira de Teses e Dissertações (BDTD) tentou utilizar essa forma de entrada de dados com as teses e dissertações, ou seja, a BDTD previa que os próprios autores registrassem e depositassem os seus trabalhos. No entanto, essa modalidade de entrada de dados não funcionou, o IBICT teve que desenvolver e implantar outra modalidade de entrada de dados, na qual as bibliotecas se encarregavam de registrar e depositar as teses e dissertações. Esta iniciativa foi adotada com o propósito de assegurar o sucesso da BDTD, o qual se denominou módulo simplificado. No Brasil, a maior parte das universidades adotou a entrada de dados centralizada por intermédio das bibliotecas e, apenas uma pequena minoria adotou a entrada de dados descentralizada, por intermédio, do autodepósito por parte dos autores.

As dificuldades encontradas pelo IBICT foram, em grande parte, devido à inexistência de qualquer autoridade sobre as instituições de ensino superior e, consequentemente isto dificultou o estabelecimento de qualquer mandato ou obrigatoriedade no processo de registro e disseminação da produção de teses e dissertações no país. Se o IBICT articulasse com as agências de fomento uma política tornando obrigatório este depósito, certamente, teria maior sucesso no registro e disseminação das teses e dissertações em meio eletrônico. Há que se

registrar que o IBICT tentou fazer tais articulações. No entanto, as negociações não avançaram devido ao que se costuma denominar a falta de interoperabilidade humana, ou seja, as articulações não prosperaram devido à falta de entendimento entre os homens. É curioso verificar, no Brasil, esse tipo de fenômeno, enquanto a interoperabilidade tecnológica é um fato, a falta de interoperabilidade humana tem se constituído em um grande empecilho para a integração e fornecimento de acesso à informação científica. Aliás, trata-se de um fenômeno muito comum no Brasil.

Nas instituições de ensino superior onde foram estabelecidos mandatos ou políticas, tornando obrigatório o registro e depósito das teses e dissertações, as suas respectivas bibliotecas evoluíram e vêm crescendo. Nas instituições onde não houve essa preocupação suas bibliotecas digitais de teses e dissertações não cresceram e continuam com a mesma quantidade de teses e dissertações registradas no seu nascedouro. No entanto, o IBICT não tem nenhuma ascendência hierárquica às instituições de ensino superior e por isso não conseguiu estabelecer a obrigatoriedade do registro e depósito das teses e dissertações.

Nos exemplos apresentados, verifica-se a necessidade de se ter uma política ou um mandato tornando obrigatório o registro da produção científica pelos próprios autores, caso contrário, a iniciativa de registro desta produção corre sério risco de fracasso. É sobre esse tema que este capítulo se propõe a discutir. Nas próximas seções serão apresentados exemplos de mandatos ou políticas que estão sendo adotadas no exterior e, em seguida, serão apresentadas as iniciativas que estão em curso no Brasil. O termo utilizado, no exterior, é mandato. No entanto, dá-se preferência ao termo política por ser mais simpático e mais abrangente. Ou seja, o termo política por ser mais abrangente propicia maiores possibilidades de disciplinar a entrada de dados em um RI.

MANDATOS OU POLÍTICAS DE ACESSO LIVRE NO MUNDO

Peter Suber apresenta em seu *blog* denominado *Open Access News (SUBER)* uma cronologia das principais iniciativas de acesso livre, que vão desde o lançamento em 1966 do Educational Research Information Center (ERIC), passando pelo surgimento da internet e das novas tecnologias da informação e da comunicação, até a aprovação pelo Congresso americano e assinatura pelo Presidente dos EUA do ato[2] que determina a todos os pesquisadores beneficiários de financiamentos de pesquisa por parte do National Institute of Health (NIH) o depósito, no repositório PubMedCentral, de seus trabalhos publicados em revistas com avaliação pelos pares (*peer review*). Neste ato é determinado que os referidos pesquisadores devem fazer o depósito dentro de um prazo de 12 meses.

O movimento do acesso livre recomenda a adoção de duas estratégias para a implantação do acesso livre ao conhecimento científico. A primeira estratégia, denominada de via dourada (*Golden Road*), refere-se às revistas científicas (ou revisadas por pares) que fornecem acesso livre aos artigos nelas publicadas. John Willinsky (2006) identificou a existência de 10 modelos de negócios que garante a sustentabilidade de revistas científicas de acesso livre. A segunda estratégia é a chamada via verde (*Green road*), a qual recomenda que os autores depositem os seus trabalhos publicados, em revistas com revisão por pares, em um repositório institucional (ou outro tipo de repositório: temático ou central) de acesso livre.

No entanto, os RI para funcionar e atender às expectativas do movimento do acesso livre dependem de algum instrumento que garanta o depósito, por parte dos pesquisadores, dos seus trabalhos publicados em revistas com revisão por pares. Em todo o mundo, as universidades vêm estabelecendo mandatos (também chamados de políticas) para garantir o povoamento dos RI.

---

2 Maiores informações sobre este ato poderão ser encontrados no sítio: http://openaccess.eprints.org/index.php?/archives/525-Protect-the-NIH-Public-Access-Mandate-From-the-Conyers-Copyright-Caricature.html, acessado em 17/02/2010.

No entanto, essas iniciativas esbarravam no embargo que as políticas e contratos assinados pelos pesquisadores ou autores junto aos editores das revistas, as quais impediam ou atrasavam o autodepósito dos trabalhos publicados nestas revistas por parte dos seus autores (pesquisadores).

Segue exemplo de mandato, que foi estabelecido pela School of Electronics and Computer Science da University of Southampton[3]:

---

3 Segue o original, em inglês, do mandato da Escola de Ciência da Computação da Universidade de Southampton:

1. It is our policy to maximise the visibility, usage and impact of our research output by maximising online access to it for all would-be users and researchers worldwide.

1a. It is also our policy to minimise the effort that each of us has to expend in order to provide open online access to our research output: http://eprints.ecs.soton.ac.uk/10688/

1b. With all our research output accessible online we will be able to respond to the research assessment and other administrative initiatives with minimal input and effort from individual staff.

2. We have accordingly adopted the policy that all research output is to be self-archived in the departmental EPrint Archive before and after peer-reviewed publication. This archive forms the official record of the Department's research publications; all publication lists required for administration or promotion will be generated from this source.

3. Our policy is compatible with publishers' copyright agreements as follows:

3a. The copyright for the unrefereed preprint resides entirely with the author before it is submitted for peer-reviewed publication, hence it can be self-archived irrespective of the copyright policy of the journal to which it is eventually submitted.

3b. The copyright for the peer-reviewed postprint will depend on the wording of the copyright agreement which the author signs with the publisher.

3c. Many publishers will allow the peer-reviewed postprint to be self-archived. The copyright transfer agreement will either specify this right explicitly or the author can inquire about it directly. If you are uncertain about the terms of your agreement, a directory of journal self-archiving policies – http://romeo.eprints.org – is available to guide you. Wherever possible, you are advised to modify your copyright agreement so that it does not disallow self-archiving.

3d. In the rare case where you have signed a very restrictive copyright transfer form in which you have agreed explicitly not to self-archive the peer-reviewed postprint, you are encouraged to self-archive, alongside your already-archived preprint, a "corrigenda" file, listing the substantive changes the user would need to make in order to turn the unrefereed preprint into the refereed postprint.

3e. Copyright agreements may state that eprints can be archived on your personal homepage. As far as publishers are concerned, the EPrint Archive is a part of the Department's infrastructure for your personal homepage.

4. We do not require you to archive the full text of books or research monographs. It is sufficient to archive the references along with the usual metadata.

5. Some journals still maintain submission policies which state that a preprint will not be considered for publication if it has been previously 'publicised' by making it accessible online. Unlike copyright transfer agreements, such policies are not a matter of law. If you have concerns about submitting an archived paper to a journal which still maintains such a restrictive submission policy, please discuss it with the Department's IPR and Copyright Advisor.

1. É nossa política maximizar a visibilidade, uso e impacto da nossa produção de pesquisa, maximizando o seu acesso *online* para todos os futuros usuários e pesquisadores do mundo todo.
    1a. É também nossa política minimizar o esforço que cada um de nós tem que gastar a fim de proporcionar o acesso livre *online* a nossa produção de pesquisa: http://eprints.ecs.soton.ac.uk/10688/.
    1b. Com toda a nossa produção de pesquisa acessível *online*, seremos capazes de responder à avaliação de pesquisa e outras iniciativas administrativas com mínima entrada de dados e esforço da equipe de funcionários.
2. Estamos de acordo com a política de que todos os resultados de pesquisa devem ser autoarquivados no repositório institucional (RI) do departamento antes e após a publicação revisada por pares (*peer-reviewed*). Este autoarquivamento no RI representa o registro oficial das publicações resultado de pesquisas do Departamento, todas as listas de publicações necessárias à administração ou a promoção será gerada a partir desta fonte (RI).
3. Nossa política é compatível com os acordos de direitos autorais, contratos ou políticas editoriais da seguinte forma:
    3a. Os direitos autorais do preprint[4] pertencem inteiramente ao autor e, portanto, pode ser autoarquivados, independentemente da política de direitos autorais da revista em que o artigo será finalmente publicado.
    3b. Os direitos autorais do artigo após a sua revisão por pares (*peer-reviewed postprint*) dependerá do teor do acordo de direitos autorais que o autor assinou com a editora da revista.
    3c. Muitos editores permitirão que um artigo (*a peer-reviewed postprint*) seja autoarquivado. O acordo de transferência de direitos autorais especificará, explicitamente, este direito, ou o autor poderá obter informações a respeito disto diretamente com o editor. Se o autor não tiver

4 *Preprint* é a versão do artigo que é submetida à revista, portanto, antes do processo de revisão por pares.

certeza sobre as condições do seu contrato, um diretório de políticas de autoarquivamento – http://romeo.eprints.org – está disponível para orientá-lo. Sempre que possível, recomenda-se ao autor alterar o seu contrato de direitos autorais para que o editor não desautorize o autoarquivamento do seu artigo.

3d. No caso raro em que o pesquisador assina um termo de transferência muito restritivo de direitos autorais, no qual ele concordou expressamente em não autoarquivar o *postprint*[5] revisado por pares (*peer-reviewed postprint*), o mesmo é estimulado a autoarquivar, junto com o *preprint*, já arquivado, um arquivo de correções, listando as principais alterações que o usuário precisa fazer para transformar o *preprint* no *postprint*.

3e. Contratos de direitos autorais podem indicar que o artigo pode ser arquivado na sua página pessoal (*homepage*). No que diz respeito aos editores, o RI é parte da infraestrutura do departamento para a sua *homepage* pessoal.

4. Nós não requeremos que você arquive o texto integral de livros ou monografias de pesquisa. É suficiente arquivar as referências juntamente com os metadados usuais (título, autor, data de publicação etc).
5. Algumas revistas ainda mantêm política de submissão estabelecendo que um preprint não será considerado para publicação, se este tiver sido previamente objeto de uma pré-impressão ou depósito em um RI, tendo em vista que ele já é de conhecimento público ou é acessível em linha. Ao contrário dos acordos de transferência de direitos de autor, tais políticas não são uma questão de direito. Se você tiver dúvidas sobre o envio de um documento previamente arquivado em um RI para uma revista científica, a qual ainda mantém política de submissão restritiva,

5 *Postprint* é o artigo publicado na revista, ou seja, o pré-print revisado por pares e publicado na revista.

por favor, consulte o departamento de IPR (Direitos de Propriedade Intelectual) ou com um Assessor/especialista em direitos autorais.

Em 2006, Stevan Harnad propôs um mandato denominado de *Immediate Deposit/Optional Access Mandato* (Mandato depósito imediato/acesso opcional) também conhecido como ID/AO, o qual tinha como objetivo ser imune aos eventuais atrasos ou embargos provocados, geralmente, pelas políticas editoriais ou contratos mantidos pelas revistas científicas (HARNAD, 2006). Este mandato é também conhecido como "duplo depósito/estratégia de distribuição". A ideia é que o autor deposite o seu trabalho, em um repositório de acesso livre, tão logo saiba que o mesmo tenha sido selecionado ou aceito para publicação em uma determinada revista com revisão por pares. Em outras palavras, esta proposta recomenda o imediato depósito, em seguida à comunicação da revista ao autor de que o seu artigo está selecionado para publicação na revista. A recomendação é que seja depositado o texto referente à versão final aceita pela revista, com revisão por pares. No caso em que haja atraso ou embargo, normalmente, provocados por políticas editoriais ou contratos estabelecidos pela referida revista que impeça esse depósito, recomenda-se ao autor o depósito dos metadados que descrevem o referido trabalho, mantendo o texto integral do trabalho publicado com acesso restrito ou fechado pelo tempo que a política ou contrato assinado com a revista estabelece. Cabendo ao autor a abertura do acesso a partir do momento em que a revista permita o acesso público ao referido trabalho. Esta proposta significa que o depósito é imediato e obrigatório e que o acesso ao documento integral (artigo publicado) é opcional.

As políticas dos editores das principais revistas científicas podem ser encontradas no sítio: http://romeo.eprints.org/. Este sítio faz parte de um projeto denominado SHERPA, onde o usuário poderá encontrar tantos as políticas dos editores científicos quanto as políticas ou mandatos das instituições de ensino e pesquisa e seu endereço é: *http://www.sherpa.ac.uk/index.html.*

POLÍTICAS/MANDATOS NO BRASIL

No Brasil, existem duas iniciativas em curso: 1) definição de políticas institucionais pelas universidades e unidades de pesquisa do Ministério de Ciência e Tecnologia como requisito para a implantação dos seus respectivos repositórios institucionais; 2) discussão por parte da Câmara dos Deputados, por intermédio de algumas de suas comissões, de um projeto que estabelece a obrigatoriedade das universidades construírem os seus repositórios institucionais e torna obrigatório aos pesquisadores dessas universidades o depósito da sua produção científica.

O IBICT vem trabalhando na sensibilização da comunidade científica quanto a importância do acesso livre ao conhecimento científico desde 2004. Uma das suas primeiras iniciativas foi o lançamento do *Manifesto brasileiro de apoio ao acesso livre ao conhecimento científico* em setembro de 2005. Este manifesto foi escrito e proposto por Hélio Kuramoto e Sely Costa e lançado, pelo IBICT, em 13 de setembro de 2005, por intermédio de uma videoconferência que interligou pesquisadores em salas montadas em Brasília, São Paulo, Campinas, Rio de Janeiro, Belo Horizonte, Fortaleza e Florianópolis. O evento, transmitido via Web, foi assistido por cerca 15 mil pessoas, segundo registros da Rede Nacional de Ensino e Pesquisa (RNP).

O manifesto é um conjunto de recomendações a todos os segmentos da comunidade cientifica para a implantação do acesso livre no País. Trata-se de um guia para a discussão e o estabelecimento de uma política de acesso livre à informação científica.

Posteriormente, em meados de abril de 2007, o IBICT articulou com o Deputado Rodrigo Rollemberg a submissão de um projeto de lei com o propósito de reforçar e dar respaldo às iniciativas de acesso livre no Brasil. Este projeto de lei foi submetido, assim, à Comissão de Ciência, Tecnologia, Comunicação e Informática (CCTCI), que o aprovou agregando duas emendas e o encaminhou à Comissão de Educação e Cultura (CEC), que também o aprovou sem nenhuma adição de emendas. Esta comissão em seguida o encaminhou à Comissão de Constituição e Justiça e Cidadania (CCJC), que atribuiu a relatoria ao

Deputado Valdenir Pereira (PSB/MT), que elaborou um parecer favorável à constitucionalidade do referido projeto de lei (BRASIL, 2009). Até o momento, hoje dia 13/03/2010, o referido projeto de lei ainda não foi colocado em pauta para votação.

Basicamente este projeto de lei propõe dois artigos: 1) estabelece a obrigatoriedade de as universidades e institutos de pesquisa públicos a construírem os seus respectivos repositórios institucionais e ao mesmo tempo torna obrigatório que os pesquisadores dessas universidades depositem uma cópia dos resultados de suas pesquisas, publicados em revistas com revisão por pares, nos referidos repositórios institucionais; 2) estabelece no seu artigo segundo que seja criada uma comissão de alto nível para a discussão e estabelecimento de política nacional de acesso livre à informação científica. (BRASIL, 2007) Este projeto de lei encontra-se, portanto, neste momento, 14/02/3010, no aguardo do agendamento de uma sessão plenária da CCJC para que seja discutida e votada. Uma vez aprovado, este projeto de lei seguirá para o Senado onde certamente seguirá os mesmos trâmites que seguiu na Câmara dos Deputados.

Outra iniciativa é a das políticas institucionais de informação que estão sendo discutidas pelas universidades brasileiras em função do desenvolvimento dos seus repositórios institucionais. Isto é consequência de ação do IBICT que distribuiu 33 servidores a universidades e unidades de pesquisa do MCT. O requisito básico era de que estas universidades construíssem seus repositórios e, ao mesmo tempo, que aprovasse internamente uma política institucional de informação de forma a garantir o povoamento dos seus repositórios por parte dos seus pesquisadores.

A título de colaboração com as universidades, o IBICT propôs que estas utilizassem como modelo a seguinte política:

Modelo de política institucional de informação

A <Universidade XXXX>/ <o Instituto XXXX> considerando a necessidade de:

- preservar a produção científica desta Instituição;
- ampliar a visibilidade da produção científica desta Instituição;

- potencializar o intercâmbio desta Instituição com outras instituições;
- acelerar o desenvolvimento de suas pesquisas;
- ampliar o acesso à sua produção científica;
- facilitar o acesso à informação científica de uma forma geral;
- otimizar a gestão de investimentos em pesquisa nesta Instituição;

reconhece a importância da implementação de ações que garantam o registro e a disseminação da produção científica desta Instituição[6]. Assim, para atender a estas necessidades, o <reitor>/<diretor> da <Universidade XXXX>/<Instituto XXX> estabelece o seguinte:

- que a <biblioteca central desta instituição> / <qualquer outra unidade da instituição> fica encarregada do desenvolvimento, implantação e manutenção do repositório institucional desta Instituição, doravante, neste documento, denominado de RI;
- que este RI desenvolvido e alimentado pela nossa comunidade científica institucional[7] será de livre acesso, tanto no contexto nacional, quanto internacional;
- que este RI deverá ter capacidade de integração com sistemas nacionais e internacionais, observando-se o uso de padrões e protocolos de integração, em especial aqueles definidos no modelo *Open Archives*.
- que a comunidade científica institucional faça publicar os artigos de sua autoria ou co-autoria, preferencialmente, em publicações periódicas científicas de acesso livre ou que façam constar em seus contratos de publicação, o depósito de artigos publicados (pos-prints[8]) em repositórios de acesso livre;
- que o aludido depósito citado no item anterior deverá ser realizado imediatamente após a comunicação de sua seleção para publicação na revista científica. Em caso de impossibilidade de depósito imediato, o autor ou co-autor terá um prazo máximo de 6 (seis) meses da data de publicação

6 Para efeito deste documento, produção científica é aquela constituída de resultados de pesquisa consolidados veiculados em veículos de comunicação científica que tenham revisão por pares.

7 Para efeito deste documento, considera-se "comunidade científica institucional" aquela constituída por seus docentes, pesquisadores, colaboradores pesquisadores ou docentes e alunos graduandos e pós-graduandos.

8 Para efeito deste documento, considera-se *pos-prints* a versão de um artigo já publicado em revistas científicas;

do referido artigo para depositá-lo no RI;

- que na impossibilidade de realização do depósito devido à cláusulas contratuais mantidas pelo autor com a(s) revista(s) onde o seu trabalho foi publicado, recomenda-se que se deposite uma cópia da versão original do trabalho, ou seja aquela versão que foi submetida à revista científica (pre--prints[9]), assim como, as alterações propostas pela revista que o publicou.
- que fica desobrigado de depósito no RI, os livros ou capítulos de livros que são publicados com fins comerciais ou que tenham restrições contratuais relativas a direitos autorais;
- que ficam desobrigados de depósito no RI, os artigos publicados em revistas científicas que estabelecem em seus contratos com os autores cláusulas que impedem o depósito de artigos publicados em suas revistas, em repositórios de acesso livre;
- que ficam desobrigados de depósito no RI os documentos cujo conteúdo integra resultados de pesquisas passíveis de serem patenteados ou de serem publicados em livros ou capítulos de livros que serão publicados com fins comerciais;
- que todos os documentos que não se enquadrarem nos itens iv, v e vi e que foram publicados em veículos de comunicação científica com revisão por pares ou que passaram, à exemplo das teses e dissertações, por avaliação de uma banca de especialistas, deverão ser depositados no RI;
- que de maneira a facilitar o povoamento do RI, a <biblioteca central desta instituição> / <qualquer outra unidade desta instituição> poderá promover o registro da produção científica desta instituição, mediante autorização dos autores da referida produção, seja efetuando a entrada de cada documento no RI ou importando os dados já registrados em outros repositórios.

Para o cumprimento desta política, esta Instituição estabelecerá mecanismos de estímulo, assim como ações de integração que possibilitem evitar duplicações de esforços. Além disso, a implementação desta política poderá suscitar a elaboração, discussão, regulamentação e estabelecimento de políticas e mecanismos específicos de forma a garantir a plena alimentação do repositório institucional

9 Para efeito deste documento, considera-se *pré-prints* a versão original de um artigo submetido a uma revista científica ainda não publicado

e, por conseguinte, a preservação da produção científica institucional.

Esta Política Institucional de Informação entrará em vigor a partir da data de sua publicação.

Cidade, <dia> de <mês> de <ano>.

Fulano de tal
Reitor da Universidade XXXX ou
Presidente ou Diretor da Instituição XXXX

Este modelo está em conformidade com a proposta de Harnad (2006) de depósito imediato e acesso opcional (ID/OA), além de não obrigar o depósito de documentos que estão sob contrato de direitos autorais restritivos que impeçam a sua livre divulgação ou que contenham informações passíveis de serem patenteadas. Portanto, é um modelo de política que está em conformidade com a legislação do direito autoral.

CONSIDERAÇÕES FINAIS

Vale ressaltar que os repositórios institucionais necessitam, inicialmente, para garantir o seu povoamento de políticas institucionais de informação. No entanto, a experiência mostra que com o tempo os próprios autores (pesquisadores) se convencem dos benefícios e da importância do depósito de sua produção científica e voluntariamente passam a ser defensores do RI. Isto acontece devido às estatísticas que são produzidas a partir do uso dos repositórios indicando a quantidade de *downloads* dos seus artigos depositados no RI. Estudos mostram que o depósito da produção científica tende a dar maior visibilidade aos autores e às universidades. O resultado disto é que as universidades que mantêm o seu repositório atualizado ganham maior poder de competitividade e os autores tendem a ser mais citados (SWAN, 2008).

O maior poder de competitividade das universidades que detêm um RI é comprovado pela melhora do seu *ranking* nas várias iniciativas de classificação das universidades, destacando-se, dentre elas o *webometric* (CSIC, 2009).

REFERÊNCIAS

BRASIL. **Projeto de Lei n. 1.120, de 21 de maio de 2007**. Dispõe sobre o processo de disseminação da produção técnicocientífica pelas instituições de ensino superior no Brasil e dá outras providências. Disponível em: < http://www.camara.gov.br/sileg/integras/461698.pdf>. Acesso em: 13 mar. 2010.

BRASIL. Comissão de Constituição e Justiça e de Cidadania. Parecer do Relator, Dep. Valtenir Pereira (PSB-MT), pela constitucionalidade, juridicidade e técnica legislativa deste, com emenda, e das emendas da CCTCI, com subemendas. **Parecer ao Projeto de Lei n. 1.120, em 29 de outubro de 2009**. Disponível em: <http://www.camara.gov.br/sileg/integras/707945.pdf>. Acesso em: 13 mar. 2010.

CSIC - CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS. **Ranking mundial de universidades em la red**. Jul. 2009. Disponível em: <http://www.webometrics. info/top6000_es.asp>. Acesso em: 14 mar. 2010.

HARNAD, Stevan. **The Immediate-Deposit/Optional-Access (ID/OA) Mandate:** rationale and model. Open Access Archivangelism. 2006. Disponível em: http://openaccess.eprints.org/index.php?/archives/71-guid.html. Acesso em: 17 fev. 2010.

KURAMOTO, Hélio; COSTA, SELY. **Manifesto brasileiro de apoio ao acesso livre à informação científica**. 2005. Disponível em: http://kuramoto.files.wordpress.com/2008/09/manifesto-sobre-o-acesso-livre-a-informacao-cientifica.pdf. Acesso em: 13 mar. 2010.

SUBER, Peter. **Timeline of the open access movement.** Disponível em: http://www.earlham.edu/~peters/fos/timeline.htm. Disponível em: <http://www.earlham.edu/~peters/fos/timeline.htm>. Acesso em: 17 fev. 2010.

SWAN, Alma. Why open access for Brazil? **Liinc em Revista**, Rio de Janeiro, v. 4, n. 2, p.158-171, set. 2008. Disponível em: <http://revista.ibict.br/liinc/index.php/liinc/article/viewFile/279/166>. Acesso em: 22 fev. 2010.

WILLINSKY, John. **The access principle:** the case for open access to research and scholarship. Massachusetts: The Mit Press, 2006. 287 p. Disponível em: <http://mitpress.mit.edu/catalog/item/default.asp?tid=10611&ttype=2>. Acesso em: 22 fev. 2010.

Consequências do movimento pelo livre acesso – *open access* – e o direito à informação científica

Marcia H. T. de Figueredo Lima

A Ciência da Informação se constituiu, a partir dos esforços de pesquisa científica e tecnológica engendrados na Segunda Guerra como uma metaciência: do grego *metá*, prefixo que significa, dentre outros, uma "reflexão crítica sobre".

Duas tradições bem marcadas disputaram veladamente – e disputam, ainda hoje – a eleição de temas, problemas, questões, objetos e metodologias de pesquisa, em torno da noção sobre o que seria a informação (constuto, artefato): se a informação privilegiada para os estudos da Ciência da Informação é (ou deveria ser) mesmo a informação científica e tecnológia ou, se os efeitos sociais da informação na melhoria da qualidade de vida política também seriam um foco de estudo. Decorrente desta questão, há um debate não enunciado sobre se o usuário dessa informação é o pesquisador (mais restritamente) ou o cidadão (mais amplamente), como a informação gera seus efeitos (é *comodity*, uma mercadoria, ou um bem público; a informação politiza e pode gerar mudança?). Há uma disputa sobre a consideração da informação como bem público ou como bem privado. Informação é poder? Informação é dinheiro?

Em torno dessas questões se detém uma comunidade de pesquisadores que delineia um universo de usuários identificados ora como os cientistas, ora como os cidadãos, ora como os empresários.

Os reflexos de sua origem vinculada ao esforço de guerra fazem da

produção científica e tecnológica um campo privilegiado (até hegemônico) nas discussões do tema. O desenho de um ciclo de informação que vai de produtores a usuários, passando pela mediação de editores científicos e dos serviços de organização e busca é inegavelmente compartilhado pelos pesquisadores da Ciência da Informação, bem como há um certo consenso sobre as diversas fases de coleta, armazenamento, tratamento, acesso e produção de novas informações, principalmente no âmbito, então, da Ciência e da Tecnologia.

Do que se disse até aqui, o importante é perceber que, do ponto de vista epistemológico, a discussão sobre o objeto da Ciência da Informação só permite nos dar certeza de que ele não é um universal – há correntes debatendo sobre a primazia dos enfoques e conteúdos. Mas, há um ponto que parece convergir para a aproximação em torno desse objeto – não toda a informação interessa à Ciência da Informação ou é passível de por ela ser abordada. O recorte sobre o qual este saber de debruça é a informação institucionalizada, aquela que sofre algum tipo de intervenção midiática – de edição, seleção, armazenamento. Quer nos detenhamos sobre a informação para a melhoria da cidadania, quer nos indaguemos sobre a informação para a produção, quer nos preocupemos com a importância do acesso à informação em Ciência e Tecnologia, os cientistas da informação geram informação sobre alguma forma de enunciação regulada em instâncias institucionais.

A proposta desse capítulo é refletir sobre a crescente importância dos princípios do *open access* – livre acesso – em benefício de um maior acesso à informação científica. O *open access* viria atender aos princípios de direito à informação, um *right to known*, é a tese.

A Ciência moderna teve suas raízes ilumininstas no *sapere aude* (atreve-te a saber, propagado por Immanuel Kant em *O que é o Esclarecimento*), que pressupunha um conhecimento público, um direito universalizado do ser humano saber e este saber laicizado, perdeu, lentamente, a característica da sacraidade, do indiscutível e do inquestionável. Como forma de entender o mundo, contrapôs-se, em um primeiro momento, ao saber sagrado, monadal e inquestionável da Religião. À publicidade (na política) dos atos do soberano como arma contra os arcanos (as arcas) do Estado – o segredo e a mentira – equivale, no

campo científico nascente, a obrigação de publicar o conhecimento e trazê-lo à luz – daí Iluminismo – do debate "público" – aqui um público douto, especializado de pares, mais restrito. Este debate tem por pressupostos o partilhamento da crença de que a narrativa de cada nova parte compreendida sobre o Universo será descrita e publicada, enfocando métodos e detalhando as fases e variáveis envolvidas, de uma maneira (pelo menos em tese) reproduzível pelos pares (o público leitor) e sempre passível de discussão.

De forma lenta, a ciência contemporânea foi gestada tal como a concebemos hoje, a partir daquela Modernidade, adquirindo sua feição atual mais marcante também induzida pelo fomento estatal decorrente do esforço de Segunda Grande Guerra nos países centrais. A partir dessa época, uma ideologia da publicação (quantitativa) se instaura e o cientista tem que publicar para ter visibilidade: é o "*publish-or-perisch*"[1]. Exemplos dessa palavra de ordem não faltam. Robert Day, em 1990, no *Boletin de la Oficina Sanitaria Panamericana* escreveu: "O cientista não só tem que fazer Ciência, mas também escrevê-la".

Neste recorte de análise privilegiado pela Ciência da Informação, o da informação científica e tecnológica, os produtores são identificados como universidades, centros de pesquisa, laboratórios de Pesquisa e Desenvolvimento (P&D), sociedades científicas, editores de periódicos científicos, serviços de indexação e resumo, livreiros, indústria de informática: *hardware* e *software*; os intermediários como os serviços de bancos de dados, portais de periódicos, *Web of Science*, bibliotecas e centros de documentação. Finalmente, os usuários são identificados com a própria comunidade científica, o que gera um ciclo no qual os produtores primeiros serão os usuários em última instância.

Este ciclo de produção, intermediação e uso (preferimos essa expressão do que consumo) de informação científica e tecnológica se organiza e gera toda uma rede econômica de interesses que envolve atores variados como os laboratórios das grandes indústrias farmacêuticas,

1 *"Publicar ou perecer"* – estudos acerca da origem da expressão remontam a uma anedota jamais publicada do Dr. Kimball C. Atwood, um geneticista da Columbia University que a teria escrito em uma carta em torno de 1950. Inquirido a respeito, ele afirmou ter escutado essa frase em um encontro de pesquisadores.

países centrais e periféricos, a poderosa indústria de tecnologias de informação, passando pela indústria editorial que movimenta alguns bilhões de dólares por ano, repetindo um tradicional ciclo que vai dos países centrais aos periféricos, sem mudanças revolucionárias no cenário geopolítico que venha a destacar qualquer revolução neste tema. A lógica de produção oligopolizada capitalista se instaurara na Ciência, como em diversos outros domínios da vida humana.

> Nos últimos 30 anos, o preço das revistas científicas tem crescido firmemente. Enre 1975 e 1995, os preços cresceram de 200 a 300% além da inflação. Isto foi acompanhado por uma queda nas assinaturas tanto por pesquisadors a título individual, como por bibliotecas, cujos orçamentos ficaram espremidos. Efetivamente, os preços das revistas ultrapassaram de longe a evolução dos orçamentos das bibliotecas, o que fez aumentar em um ritmo um pouco inferior o total dos orçamentos de pesquisa acadêmica (DEWATRIPONT, 2006, p. 8, tradução nossa).[2]

Se anuncia a partir da década de 1970, com maior clareza aparecendo nos anos 1980, segundo Mueller (2006), uma crise nesse modelo de publicação, até porque, os cientistas de países periféricos já se ressentiam desse modelo oligopolizado, que não podia mais ser sustentato nem mesmo pelas bibliotecas americanas. Como reação, a partir dos anos 1990, vemos lentamente delinear-se e crescer uma nova forma de disponibilização do saber produzido pela Ciência, possibilitada em larga escala pela Internet. Não devemos esquecer que a rede nasceu nos centros universitários (ainda que tenha sido tomada pelo *e-commerce*). Esse novo modo de publicar na Ciência trata-se do movimento pelo livre acesso. Este movimento cresce, e tem sua culminância institucional ao fim da primeira década do segundo milênio. Uma espécie de reação a esse modelo brevemente mencionado de grandes empresas

2 "In the last 30 years, the prices of scientific journals have been steadily increasing. Between 1975 and 1995, they increased 200%-300% beyond inflation. This was accompanied by a fall in subscriptions both by individual researchers and by libraries whose budgets got squeezed. Indeed, journal prices far outpaced the evolution of library budgets, which did increase at a somewhat slower pace than total academic research budgets."

– os editores científicos – oligopolizando a publicação na Ciência. Essa reação se consubstancia em um conjunto de padrões e ideais em torno da ideia de arquivos abertos, mas conhecidos pela sigla da expressão em inglês OA (*Open Archives*).

Esse movimento cresce, então, não por acaso, mas como consequência e reação a esse modelo oligopolizado de produção científica, até assumir a feição de política pública institucionalizada nos termos do Direito dos Tratados. Em 25 de março de 2007, no transcurso das comemorações do 50º aniversário do *Tratado de Roma* que criou a União Europeia, foi assinado o *Tratado de Berlin* que aborda o *Open Access* ou Acesso Livre ao Conhecimento nas Ciências e nas Humanidades. No próprio texto da Declaração, há um enunciado afirmando que suas origens se encontram na *Declaration of the Budapest Open Acess Initiative*, fruto do encontro realizado em 01 e 02 de dezembro de 2001 da ECHO Charter e da *Bethesda Statement on Open Access Publishing*, preconizando um modelo de acesso democrático a conteúdos e ferramentas *(softwares)* como forma de viabilizar a universalização do conhecimento.

As iniciativas para a culminância da assinatura desse tratado, no entanto, se gestaram processual e lentamente. O contexto, segundo Kuramoto e Costa (2005), vinha sendo acompanhado desde o ano de 2000. Em 13 de setembro de 2005, o Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT) lança o *Manifesto brasileiro de apoio ao acesso livre à informação científica.*

Em artigo da *Nature*, Steve Lawrence (2001) afirmou que "artigos on line disponibilizados gratuitamente são mais frequentmente citados. Para maior impacto e mais rápido progresso científico, autores e editores deveriam prover uma pesquisa de fácil acesso".

De forma bastante esquemática, podemos resumir essa forma de publicação como aquela em que o autor disponibiliza seu artigo na Internet, adiando a fase de *peer review* – a revisão pelos pares que ocorria – na forma de publicação em papel - *ex ante*, antes do conhecimento do publico científico, por um modo de apropriação e revisão mais difusa e *ex-post*. O artigo publicado em papel era revisado e criticado por um grupo de *experts*, em geral três *referees* que opinavam sobre a qualidade do artigo em termos de novidade e coerência com respeito a

certas regras bem como o interesse para a revista, o grau de novidade científica, o grau de contribuição para o saber partilhado pela comunidade no momento da submissão do artitgo ao editor do periódico. As publicações científicas que optaram por disponibilizá-las em formato eletrônico continuam, inclusive, atuando pelo modo da revisão pelos pares; o diferencial está na maior agilidade de transação entre o "manuscrito" em formato digital que circula mais rápido entre autores, editores e revisores. Sem dúvida, mesmo a forma de publicação eletrônica com intermediação editorial mantém o esquema de revisão pelos pares *ex ante.* Mas, na forma de publicação preconizada pelos princípios do *open access*, essa forma de revisão *ex ante* é substituída pela forma de revisão *ex post*. O autor utiliza repositórios institucionais ou outros especializados para disponibilizar resumos ou textos completos de seus trabalhos. Se por um lado, não sofrerá restriçõs de *referees*, haverá sempre um grupo de voluntários que auxiliam na crítica dos seus originais. Continua de outra forma, o *peer review.*

O que, brevemente mencionamos acima, já dá conta de uma das características dessa forma de publicação: a exposição pública e ágil das conclusões, uma forma de revisão posterior (e, nem por isso, menos especializada do que aquela dos *referees*) e, principalmente, uma quebra do esquema monopolista de publicação dos grandes editores científicos.

Assim, no ciclo da produção e uso de conhecimento científico, vemos instaurarem-se novos atores: periódicos eletrônicos, repositórios de documentos digitais, bibliotecas digitais, portais temáticos, provedores de serviços, motores de busca especializados em Ciência, índices de citações autônomos, diretórios de repositórios de documentos digitais, *blogs* científicos.

Vemos, também, instaurar-se uma nova lógica, a da publicação pelo próprio autor, sem suporte financeiro e sem retorno através das subscrições de assinaturas (que eram apropriadas por um editor), um sistema de cooperação voluntária de revisão – que, diga-se, não é de todo nova, já que os *referees* dos periódicos em papel também não eram remunerados financeiramente, circulando a moeda do prestígio acadêmico e, principalmente, uma maior agilidade no processo de publicação – chegada do texto completo ao público leitor. Uma outra

consequência importante, decorrente dessa última é uma diminuição relativa da importância do periódico científico tradicional (editado comercialmente) na comunicação da Ciência.

Ao contrário da remuneração dos editores pelos assinantes das revistas, há alguns sistemas de publicação que cobram dos autores o custo da revisão pelos pares (o sistema *pay-per-publish*). Essa é outra mudança importante: no sistema tradicional, os leitores pagavam o custo de publicação. Nessas novas formas de publicação, os autores se organizam e pagam em troca do reconhecimento e validação de suas propostas acadêmicas.

Como se pode deprender da leitura até aqui, a efetivação de esforços em torno do *Open Access* engendrou consequências de natureza filosófica que aproxima cada vez mais o autor da realização dos ideais iluministas, de um saber amplamente divulgado.

Já que o objetivo desse artigo não é apresentar todas as nuances dessa nova forma de relação entre esses atores, interessa-nos refletir sobre o que isso representa em termos de direito à informação científica.

O DIREITO À INFORMAÇÃO CIENTÍFICA

Costuma-se definir o ramo do Direito à Informação como um conjunto de saberes jurídicos em torno do objeto informacional. Este novo ramo do Direito tem aparecido (no Brasil) em cursos *lato sensu* e tematiza questões em torno do direito autoral, da propriedade industrial, o direito *sui generis* (direito de propriedade dos proprietários de bases de dados). Nesses cursos, pouco se vê sobre uma informação alargada para a cidadania ou a temática da maior transparência governamental.

Patrícia Reyes Olmedo (2004) cita autores por demais conhecidos nesta área e sinaliza o ano de 1948 como marco inicial de emergência deste tema na contemporaneidade. Alguns autores sinalizam que houve um encontro anterior à Assembleia da ONU, na mesma cidade de Genebra, no mesmo ano, sobre a questão. Após estes dois enclaves, sinalizam-se, ainda, na sequência histórica a Convenção Americana e a Convenção Europeia acerca desse direito.

Desantes Guanter (1990), professor na Espanha e diversos países latino-americanos, desta disciplina, e Zaffore (2000) citados por Reyes Olmedo (2004) afirmam que

> o direito à informação, como direito humano subjetivo, deu lugar à consideração científica de todas as normas, mais ou menos dispersas, que regulavam as atividades informativas e as mensagens que com respeito a elas era possível difundir. Estas normas, unificadas pelo critério de serviço à realização do direito à informação, constituem o ordenamento jurídico informativo, denominado Direito da Informação, e que segundo Zaffore constitui 'o saber jurídico que se ocupa de estudar, sistematizar e ordenar os instrumentos e instituições jurídicas que corporificam e regulam o direito à informação'.

Reyes Olmedo (2004) sinaliza, ainda, que o grande objetivo do sistema legal formado por esse subconjunto de dispositivos legais, concernentes ao tema informacional e discutidos neste novo ramo disciplinar, tem em vista comprovar a eficácia desse direito humano contemplado legalmente pelos institutos que dele fazem parte: "a validade ou legitimidade do Direito da Informação se encontra em que favoreça à realização do direito à informação". Quer dizer: o **Direito da/de** Informação, enquanto novo campo disciplinar, irá retirar sua legitimidade do **direito à** informação, que é o seu fundamento principal. E o objeto que dá autonomia epistêmica a esse saber é o mesmo objeto da Ciência da Informação: a informação.

Segundo esse raciocínio, a visão do objeto informacional autonomizado reveste esse ramo de uma concepção individualista, porque proveniente dos direitos individuais. Uma perspectiva contrária, que parece mais ampla, é o **caráter coletivo do direito de acesso à informação pública**, que considera a transparência e a publicidade como bem público e social e, neste sentido, um mecanismo efetivo de controle das instituições, base legitimadora do exercício do poder (REYES OLMEDO, 2004).

Tomado neste aspecto, o direito à informação, assim autonomizado, desvinculado do caráter acessório para atingir outros bens, assume o caráter de basilar para a democracia, como um elemento para alcançar a real e efetiva participação cidadã.

Neste marco teórico, nos termos de um direito comparado à informação, que as legislações nacionais ainda têm muitos aspectos em que avançar, Reyes Olmedo (2004) concorda com nossa visão ao afirmar que "quando falamos de direito de acesso à informação pública, não falamos de outra coisa, que não seja o direito que têm os cidadãos de tomar conhecimento sobre os feitos, atos e documentos emanados do Estado".

Para Celso Lafer (1998, p. 241 e 25)

> O direito à informação, que no Direito das Gentes[3], como o direito à intimidade, tem como objeto a integridade moral do ser humano, é precipuamente *uma liberdade democrática, destinada a permitir uma adequada, autônoma e igualitária participação dos indivíduos na esfera pública*. [...] [O poder] sempre resulta do agir conjunto, que se baseia no direito de associação e que requer a comunicação entre as pessoas e, portanto, o direito à informação.

Para Hannah Arendt o poder legítimo se funda na adesão política (decorrente do direito de reunião) dos muitos à norma, já que legitimidade é entendida como adesão ao comando legal. Poder e legitimidade são fenômenos plurais e coletivos distintos da violência.

---

3 "Direito das gentes" – é a denominação que os antigos escritores davam ao Direito Internacional Público, sendo assim compreendido como o conjunto de regras que regem a vida das nações entre si. (DE PLÁCIDO E SILVA, 2002. p. 271). Segundo Hespanha (1994, p. 91), no *Digesto*, podemos ter notícia de quão remota é a noção: o jurista romano Gaio assim se pronunciava: "todos os povos que se regem por leis e costumes usam um direito que em parte é seu próprio, em parte comum a todos os homens. E assim aquele direito que cada povo institui para si chama-se próprio da cidade [...]. No entanto, aquele a que a razão natural institui entre todos os homens e que entre todos é observado chama-se direito das gentes [...]". Mesmo mais recentemente, o professor, ex-chanceler e atual juiz da Corte Internacional de Haia, Francisco Rezek inicia a 9ª edição de seu livro, lançado pela Saraiva em 2002, *Direito Internacional Público*: curso elementar" à página 11 com uma referência ao "*direito das gentes*". Ele inicia, assim, o título: *O Tratado Internacional – Perspectiva histórica: parte fundamental do direito das gentes, o direito dos tratados ...*".

Ao mesmo tempo em que um saber sobre as coisas públicas da cidade se configura como tema de um direito à informação, é comum falar-se em domínios regionalizados de um direito à informação ambiental, ou um direito à informação consumerista (informação ao consumidor) e podemos demarcar um direito à informação científica. Nesses domínios regionalizados, há também um direito de saber.

Assim, a reunião de muitos em torno de um objeto coletivo é uma forma política de abordagem ou forma de lidar com um evento. – o *Open Access* seria a realização de um modo coletivo e, portanto, político de lidar com o acesso à informação científica. Sem dúvida uma das formas de acesso à informação que, aparentemente, agiliza a trajetória da informação da geração ao uso pela comunidade científica.

Em termos de avaliações tendenciais para o futuro poderíamos antever um crescente movimento coletivo em torno do direito ao saber: o *Open Access* à informação científica permitirá com certeza uma maior *right to know*. Sobre a magnitude dos impactos que isso trará para o volume e a agilidade da informação, talvez seja possível compará-la ao impacto que a prensa representou sobre o documento manuscrito: um aumento da esfera dos leitores de então comparável a uma explosão na agilidade de acesso pela comunidade científica aos resultados das pesquisas.

Há, sem dúvida, outras consequências, como o aumento do número de bases de texto completo, a necessidade de sistemas cada vez mais interoperáveis, em detrimento dos sistemas fechados representados pela publicação em papel, um impacto global mais difuso (permitido pela presença em larga escala de instumentos computacionais associados como os tradudores automáticos), o aumento da visibilidade dos autores que publicam em rede como forma legitimadora de seu saber e conquista de prestígio acadêmico, além da facilidade imcomparável de publicação e diminuição de custos (a diminuição do uso de papel e, mesmo, a pulverização da impressão pelo usuário final).

Os próprios tradicionais editores científicos também estão alterando suas práticas: já se tem notícia de que alguns aceitam para publicação em papel artigos já publicados na rede, os quais são considerados como *pre-prints*. Outro sinal de abertura dos editores é a crescente

prática de permissão de livre acesso a fascículos anteriores.

Exercícios de futurologia não são bem vistos na comunidade acadêmica e essas novas formas de publicação mais abertas, menos reféns da interferência comercial, mais cooperativas são demasiado novas para que apreendamos todas suas consequêcias. Nesse sentido, convém ao pesquisador observar com cautela os rumos que essas novas práticas de escritura e edição hão de tomar.

REFERÊNCIAS

DAY, Robert A. Como escribir y publicar trabajos cientificos. **Boletin de la Oficina Sanitaria Panamericana**, v. 109, n. 1, p. 46-59; n. 2, p. 152-171; n. 3, p. 267-288, 1990.

DESANTES, José María. **Información y derecho**. Santiago: Pontificia Universidad Católica de Chile, 1990. (Colección Actualidad e Información). Apud REIS OLMEDO, Patrícia. **El derecho de acceso a la información pública**: desafios de un mundo globalizado y democrático. In: CONGRESSO INTERNACIONAL DE DIREITO E TECNOLOGIAS DE INFORMAÇÃO, 3., 2004. Salvador: Instituo Brasileiro de Direito de Informática, 2004.

DEWATRIPONT, M. et al. **Study on the economic and technical evolution of scientific publications market in Europe**. Final Report – 2006. Brussels: DG-Research, European Commission, 2006. Disponível em: <http://dipot.ulb.ac.be:8080/dspace/bitstream/2013/9545/1/md-0005.pdf>. Acesso em: 09 out. 2009.

DE PLÁCIDO E SILVA, Oscar Joseph. **Vocabulário jurídico**. 20. ed. Rio de Janeiro: Forense, 2002. p. 271.

HESPANHA, Manuel. A ordem jurídica do antigo regime. In: ______. **História de Portugal moderno político e institucional**. Lisboa: Universidade Aberta, 1994. p. 87-93.

KURAMOTO, Hélio; COSTA, Sely. **Manifesto brasileiro de apoio ao acesso livre à informação científica. Disponível em: <**http://kuramoto.blog.br/2006/03/12/o-manifesto-brasileiro-de-apoio-ao-acesso-livre-a/>. Acesso em: 09 out. 2009.

LAFER, Celso. **A reconstrução dos direitos humanos**: um diálogo com o pensamento de Hannah Arendt. São Paulo: Companhia das Letras, 1998.

LAWRENCE, Steve. On line or invisible? **Nature**, v. 411, n. 6837, p. 521, 2001.

MUELLER, Suzana Pinheiro Machado. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**, Brasília, v. 35, n. 2, p. 27-38, maio/ago. 2006.

REIS OLMEDO, Patrícia. **El derecho de acceso a la información pública**: desafios de un mundo globalizado y democrático. In: CONGRESSO INTERNACIONAL DE DIREITO E TECNOLOGIAS DE INFORMAÇÃO, 3., 2004. Salvador: Instituo Brasileiro de Direito de Informática, 2004.

REZEK, Francisco. **Direito internacional público**: curso elementar. 9. ed. rev. São Paulo: Saraiva, 2002.

ZAFFORE, Jorge. **Información social**: Derecho y Regulación. Buenos Aires: Ediciones Depalma 2000. Apud REIS OLMEDO, Patrícia. **El derecho de acceso a la información pública**: desafios de un mundo globalizado y democrático. In: CONGRESSO INTERNACIONAL DE DIREITO E TECNOLOGIAS DE INFORMAÇÃO, 3., 2004. Salvador: Instituo Brasileiro de Direito de Informática, 2004.

Implantação do repositório institucional da Universidade Federal da Bahia:

Uma política de acesso à produção científica[1]

Flávia Garcia Rosa

INTRODUÇÃO

Em 2007, com o objetivo de disponibilizar os conteúdos da produção da Editora da Universidade Federal da Bahia (EDUFBA) em acesso livre, buscou-se o apoio do Centro de Processamento de Dados da UFBA (CPD) para propor a Universidade a criação de um Repositório Institucional (RI) a partir da instalação do *software* livre, Dspace.

Essa proposta originou-se de uma das recomendações da dissertação de mestrado *Pasta do professor: uso de cópias nas universidades,* defendida em outubro de 2006 no Programa de Pós-Graduação do Instituto de Ciência da Informação (ICI) da UFBA. O objetivo da proposta era minimizar o uso de cópias de livros dando acesso livre à produção da Editora.

Havia limitações por conta do CPD com relação ao conhecimento de implantação de repositórios e a Editora da UFBA disponibilizou artigos científicos que tratavam do tema, bem como o próprio CPD pesquisou o Dspace – *software* de código aberto e acesso livre de suporte ao desenvolvimento de RI. Além disto, contou-se, desde o início, com

1 Esse capítulo faz parte da tese intitulada *A disseminação da comunicação científica da Universidade Federal da Bahia*: acompanhamento e avaliação crítica da implantação do Repositório Institucional em desenvolvimento no Programa de Pós-Graduação Multidisciplinar em Cultura e Sociedade da Faculdade de Comunicação da UFBA. Tem como orientador o Professor Marcos Palácios da Facom/UFBA e no estágio doutoral na Universidade do Minho, a Professora Maria João Gomes. A parte técnica deste texto contou com a colaboração de Rodrigo Meirelles. Bolsa sanduiche financiada pela Capes (29 de novembro de 2009 a 3 de abril de 2010).

o suporte tecnológico de Rodrigo Meirelles, bibliotecário e mestre em Ciência da Informação pelo ICI\UFBA, que já dominando o Sistema de Editoração Eletrônica de Revistas (SEER), se interessou em participar voluntariamente do projeto e se aprofundou na implantação do *software* escolhido participando das discussões em torno da gestão do repositório.

Ainda em setembro de 2007, o Dspace foi instalado no servidor da UFBA, e percebeu-se, então que para alcançar os resultados desejados, a criação do RI deveria ser repensada na forma de um projeto mais amplo e não restrito a EDUFBA.

Essa implantação do RI foi então transformada em projeto de pesquisa de doutorado, submetida à seleção e aprovada para o ano letivo de 2008, no Programa de Pós-Graduação Multidisciplinar em Cultura e Sociedade da Faculdade de Comunicação. Para a estruturação do projeto, pensando o RI como uma política da Instituição, as principais motivações foram: a possibilidade de redução do uso de cópias de livros no ambiente acadêmico; ampliação da visibilidade da pesquisa da UFBA em todas as áreas, possibilitando que a área artes (Dança, Música, Teatro e Artes Plásticas) disponibilizem conteúdos em outros formatos que não apenas textuais; e consequentemente, contribuir de certa forma com o desenvolvimento da ciência do país e com a redução da exclusão cognitiva.

Quanto aos seus objetivos, é possível classificar essa pesquisa como de nível descritivo-exploratório, que utiliza como método de investigação a pesquisa-ação. Como tal, necessita de um planejamento de trabalho, obedecendo um cronograma de comum acordo com os envolvidos na área da pesquisa e definindo um plano de ação e resultados esperados. (GIL, 1999; SPINK, 1976; THIOLLENT, 2005). Essa metodologia é de grande valia em investigações cuja intenção é desenvolver o conhecimento por meio da interação entre pesquisador e o elemento da pesquisa, estabelecendo uma relação ativa entre o pesquisador e os envolvidos na área pesquisada. (SPINK, 1976)

Adota-se como definição para Repositório Institucional para essa pesquisa o que diz Clifford Lynch (2003, p. 2, tradução nossa), diretor da União para Informação em Rede:

> [Repositório Institucional é] Um conjunto de serviços que a universidade oferece para os membros de sua comunidade para o gerenciamento e a disseminação de conteúdos digitais, criados pela instituição e membros da sua comunidade. É essencialmente um compromisso organizacional com a gestão, desses conteúdos digitais, inclusive preservação de longo prazo, quando apropriado, bem como organização e acesso ou distribuição.[2]

O RepositoriUM da Universidade do Minho (UM) em, Braga, Portugal, primeira instituição de língua portuguesa a instalar um RI, está servindo de modelo para esse projeto da UFBA. A implantação desse Repositório iniciou-se em maio de 2003, e contou com uma importante etapa que foi a tradução da interface gráfica para a língua portuguesa, contribuindo para que os outros países de mesma língua já não precisassem cumprir essa etapa. No final do primeiro ano, podia se afirmar sobre o RI "[...] que, genericamente, o projecto [era] um sucesso e que o conceito de repositório institucional [estava] na Universidade do Minho para durar." (RODRIGUES et al., 2004)

Em junho de 2008, em visita técnica a esta Universidade foi possível dialogar com o Prof. Dr. Eloy Rodrigues, diretor do Centro de Documentação e idealizador do projeto de implantação do RI dessa Instituição. Na ocasião, a principal questão levantada por Rodrigues foi a ausência de uma política institucional quando ocorreu a implantação do RI na Universidade da Minho, questão que foi corrigida e com isso se ampliou a adesão da comunidade ao autoarquivamento permitido por essa modalidade de arquivo aberto e do *software* utilizado.

Ainda em junho, antes da visita a Universidade do Minho, o projeto foi apresentado ao Reitor da UFBA, Prof. Dr. Naomar de Almeida Filho, que, de imediato, entendeu que o RI deveria, de fato, ser encaminhado como uma política da Instituição.

---

2 "a set of services that a university offers to the members of its community for the management and dissemination of digital materials created by the institution and its community members. It is most essentially an organizational commitment to the stewardship of these digital materials, including long-term preservation where appropriate, as well as organization and access or distribution."

Esse texto tem como objetivo, relatar a experiência da Universidade Federal da Bahia no processo de implantação do seu Repositório Institucional, ainda em andamento bem como apresentar o contexto histórico da política de disseminação da informação na Instituição.

CONTEXTO HISTÓRICO

A Universidade da Bahia foi criada em abril de 1946, pelo Decreto-Lei nº 9.155, assinado pelo Presidente da República Eurico Gaspar Dutra e pelo Ministro da Educação Ernesto Souza Campos. Em julho do mesmo ano, reunido para eleger o Reitor, o Conselho Universitário, por votação unânime, indicou o Professor Edgard do Rêgo Santos para ocupar o cargo. A posse ocorreu no mesmo ano, em cerimônia solene no dia 2 de julho, data magna da Bahia (UFBA, 1967).

Para instituir a Universidade da Bahia, inicialmente reuniu-se a centenária Faculdade de Medicina – fundada em 1808, primeira Escola de Medicina da Colônia, com os cursos de Cirurgia Anatômica e Medicina – a Faculdade de Filosofia da Bahia, as escolas de Odontologia e Farmácia que funcionavam juntamente com Medicina e a Faculdade de Ciências Econômicas. (UFBA, 1971) Com a promulgação da Lei nº 1.254 de 1950 instituiu-se o Sistema Federal de Ensino Superior e a Universidade da Bahia passou a ser denominada Universidade Federal da Bahia (UFBA).

O reitorado de Edgard Santos estendeu-se por um período de 15 anos. Em 1961 foi destituído do cargo e suas ações destacaram-se pelo incentivo ao ensino universitário, de modo geral, mas, principalmente, pela implantação de estudos superiores em áreas que, na época, eram consideradas "supérfluas". "Propostas ousadas e experimentos inéditos na arte e culturas brasileiras tiveram o seu lugar nesses anos patrocinados, de certa forma pela Universidade Federal da Bahia, sob o comando do Reitor Edgard Santos" (FERRAZ, 1995, p. 8). Através da criação de escolas de artes e de institutos de extensão cultural, a Universidade passou a participar dos movimentos culturais da Bahia. Em 1955 foi criado o Seminário de Música e, no ano seguinte, as escolas

de Teatro e Dança. Em 1959 foi inaugurado o Museu de Arte Sacra.

Embora haja uma tendência de se destacar a gestão de Santos pelos movimentos renovadores nas artes, literatura e humanidades, acrescenta-se a esses, uma participação ativa com a inserção da universidade em ações de planejamento do desenvolvimento científico do estado da Bahia. (DIAS, 2005) A criação do curso de Geologia (1957) com o apoio da Petrobrás, empresa com presença marcante na Bahia, do curso de Administração, que tinha como objetivo atualizar as formas de gerenciamento e do Laboratório de Geomorfologia e Urbanismo, sob o comando de Milton Santos são alguns exemplos. A extensão universitária configurou-se com a criação do Centro de Estudos Afro-Orientais (CEAO) idealizado pelo humanista português Agostinho Silva (RUBIM, 1999).

Quanto à questão da disseminação da produção científica na UFBA no reitorado de Edgar Santos, houve como iniciativa um programa editorial: *Publicações da Universidade da Bahia*, responsável pela "[...] difusão da cultura, [...] edição de trabalhos originais, científicos e literários, nacionais e estrangeiros e [...] de autores baianos" (UNIVERSIDADE, 1967, p. 72).

As primeiras publicações nas quais constavam em suas capas o nome da Universidade Federal da Bahia foram aquelas resultantes de coedições entre a universidade e a Livraria Progresso Editora, de Manoel Pinto de Aguiar. Em entrevista concedida à poeta Myriam Fraga, Pinto de Aguiar declara que:

> [...] em 1957, o reitor Edgard Santos, desejando ampliar o raio de atuação da Universidade, incubiu-me das publicações, firmamos um convênio pelo qual fazíamos edições conjuntas, ficando parte para a Universidade e parte para a livraria para distribuição comercial [...] (A AVENTURA..., 1993, p. 108).

Em 1959 a Universidade Federal da Bahia inicia suas atividades editoriais. Após 13 anos de sua fundação se criou informalmente o Departamento Cultural, ligado à Coordenação Central de Extensão e responsável pela publicação do *Jornal da Universidade*, da revista

Universitas - Revista de cultura da Universidade da Federal da Bahia, além de outras coleções de cunho cultural (ROSA, 1994).

Com a reforma universitária em 1968, o Departamento Cultural foi extinto, surgindo em março de 1970, o Programa de Textos Didáticos – idealizado pelo Reitor Roberto Santos – responsável pela publicação de textos encaminhados pelos diversos professores\pesquisadores. Dentre eles tiveram seus originais publicados nesta fase: Luis Henrique Dias Tavares, Maria de Azevedo Brandão, A. L. Machado Neto, Inaiá Carvalho, Romano Galeffi, Délio Pinheiro e István Jancsó, hoje professor da USP, dentre outros. Publicava-se nesta fase uma média de dois títulos por mês.

Em 1971, foi formalmente criado o Centro Editorial e Didático (CED) da UFBA para desenvolver as atividades do extinto Departamento Cultural, englobando o programa de Texto Didático e o Núcleo de Recursos Audiovisuais, que passou a funcionar em 1972 com esta configuração e organizando a atividade editorial da Universidade. Em 1974, o CED passa a ocupar o endereço onde hoje funciona a Editora da UFBA, no Campus da Federação\Ondina, reunindo a partir daí os núcleos de Publicações e o de Recursos Didáticos, além da pequena gráfica universitária.

DISSEMINAÇÃO DA PRODUÇÃO CIENTÍFICA DA UFBA

Inicialmente não havia uma política editorial estabelecida na UFBA, o que ocorreu a partir de 1977 quando de fato o CED passa a desempenhar o papel de publicar e disseminar a produção científica e cultural da Universidade. A filosofia destas publicações era definida por um conselho editorial, criado durante o reitorado do Professor Luiz Fernando Seixas de Macedo Costa (1979 - 1983).

Em março de 1992, o Conselho universitário aprovou a proposta de transformação do CED em Editora Universitária, convertendo-o em órgão suplementar, ligado diretamente ao gabinete do reitor. Somente em 26 abril de 1993 o Conselho Federal de Educação aprovou de fato

essa transformação, já que previa a criação de novos cargos.

Em setembro de 1992, na gestão de Eliane Azevedo (1992 - outubro de 1993) e posteriormente dando seguimento ao reitorado de Azevedo, no reitorado do Professor Felippe Serpa (1994 – 1998) houve a adoção de uma política de continuidade a um trabalho de profissionalização do setor de editoração que já vinha sendo desenvolvido no reitorado do Professor Rogério da Costa Vargas (1988 – 1992) e a definição junto ao Conselho Editorial, indicado por Serpa, da política para as publicações da Universidade, ampliando e consolidado a atuação da Editora Universitária no âmbito da UFBA e junto à comunidade em geral.

Destaca-se ainda nesse período a redação do documento *Subsídios para uma política de comunicação da UFBA*, produzido por um grupo de trabalho composto por Antonio Albino Canelas Rubim (diretor da Faculdade de Comunicação), Antonio Brito (Presidente da Câmara de Extensão), Antonio Dias Nascimento (Assessor de imprensa da UFBA e presidente da comissão), Armindo Jorge Bião (Pró-Reitor de Extensão), Gustavo Fálcon (Diretor da editora da UFBA), José Benjamim Picado (Chefe de Departamento de Comunicação) e Vera Martins (Chefe de Departamento de Jornalismo). Esse foi um primeiro documento produzido no âmbito da Instituição onde consta uma política para a comunicação e trata de questões específicas da Editora. Destacamos do documento os seguintes itens:

> [...] reivindicação de apoio da instituição universitária, com linhas de financiamento para as revistas de Departamentos, Programas de Pós-Graduação, Órgãos Suplementares etc. e a manutenção de coleções de livros torna-se não só justa, mas um componente inerente de uma universidade de qualidade.
> Separar funcionalmente as atividades da editora e da gráfica, buscando a melhoria da qualidade, a divisão dos trabalhos e das responsabilidades;
> Fazer valer, já no exercício de 94, o regimento aprovado pelo Conselho Universitário e pelo CFE para proporcionar o aparato administrativo indispensável ao funcionamento profissionalizado da Editora e da Gráfica;

> Modernizar e reaparelhar as livrarias da Editora, adequando-as às necessidades de marketing atual e criar emergencialmente um posto de vendas no Pelourinho, destinado a comercialização de ensaios e outras séries universitárias não necessariamente acadêmicas;
> Montar esquema de vendas em todos os campi, com a colaboração dos livreiros independentes que atuam em quase todas as unidades da UFBA;
> Criação de um Programa universitário de apoio às revistas acadêmicas (NASCIMENTO, 1993).

O documento foi aprovado em seção do Conselho Universitário de 17 de novembro de 1994, por maioria, com algumas alterações propostas em plenário. No que se referia à Editora, algumas ações foram concretizadas como a separação da Editora e da Gráfica, abertura de ponto de vendas no Pelourinho e a manutenção de um programa editorial inclusive com coedições com outras editoras.

No reitorado de Heonir Rocha (1998 – 2002) destaca-se a criação da Portaria n. 332, de 8 de julho de 2002, que tinha como objetivo, preservar a memória da Instituição e assegurar a disseminação da sua produção acadêmica. Estabeleceu a Biblioteca Central reitor Macedo Costa/Seção Memória da UFBA, como depósito obrigatório de toda produção científica da Universidade, assim como das obras significativas das atividades acadêmica, cultural e técnica da Instituição, editadas e coeditadas pela a Editora da UFBA e pelas Unidades de Ensino, órgãos e serviços. (VIEIRA, 2000)

Em 2002 assume por um período de quatro anos o Professor Naomar de Almeida Filho, que cumpre até esta data, sua segunda gestão. No início do seu primeiro mandato foi constituído um grupo de trabalho para definir uma política para o sistema de comunicação da UFBA. O grupo, designado pelo Reitor foi composto por: Antonio Albino Canelas Rubim (diretor da FACOM e coordenador do grupo), Cláudio Cardoso (Chefe da Assessoria de Comunicação), Flávia Garcia Rosa (Diretora da EDUFBA) e dos seguintes professores da FACOM: Fernando Conceição, Linda Rubim, Paulo Leandro, Simone Bortoliero e Washington Souza Filho.

As propostas desse grupo de trabalho para a EDUFBA foram levadas para discussão no Conselho Deliberativo da Editora e a política foi adotada como norteadora para as ações. O documento tratava, em cinco itens, de metas estratégicas, modo de organização, propostas de atividades, propostas de coleções – foram criadas quatro coleções: Manuais, Bahia de Todos, Sala de Aula e Pré-universitária – e propostas para divulgação.

Em 2005 foi criada através da Portaria nº 080 de 8 de março de 2005, a Biblioteca Digital da UFBA, tendo como objetivo "assegurar a disseminação da sua produção acadêmica". (UFBA, 2005)

A produção científica da UFBA vem se ampliando a cada ano, consequencia natural da ampliação do número de Programas de Pós-graduação e grupos de pesquisa, no entanto a visibilidade dessa produção não corresponde a esse crescimento. A posição ocupada pela UFBA no Ranking Web das Universidades do Mundo[3], elaborado pelo Centro Nacional de Pesquisa da Espanha, em junho de 2009, era a 422ª posição, a 19ª no continente Latino Americano e a 12ª posição em relação às demais universidades brasileiras. Em janeiro 2010 essas posições se alteram para os seguintes índices: 555ª posição com relação as universidades do mundo, a 16ª no continente Latino Americano e a 11ª posição em relação às demais universidades brasileiras. Há, no entanto, um descompasso evidente entre a primeira colocado no Brasil e a UFBA, sobretudo no item visibilidade e produção acadêmica disponível e citada. A primeira colocada em termos de visibilidade ocupa a 61º posição e produção científica disponível e citada 12ª, enquanto a UFBA ocupa a 669ª e 133ª posições respectivamente, embora em relação a julho de 2009, a UFBA tenha melhorado sua posição em relação às instituições da América Latina e ao Brasil.

A necessidade de ampliar a visibilidade da produção científica da universidade repercute tanto para a sua comunidade interna que passará a acompanhar o que está sendo produzido na Instituição e poder acessar facilmente e por outro lado, em inserir a Instituição na produção científica nacional e mundial, graças às ferramentas disponíveis. Segundo Sanchez Tarragó (2007)

---

3 http://www.webometrics.info/

> Os países subdesenvolvidos dependem do uso intenso dos resultados das atividades científicas e técnicas para fazer avançar suas sociedades. Sem dúvida, o abismo existente entre esses países e os desenvolvidos com respeito ao acesso, criação e utilização dos conhecimentos científicos se aprofunda cada vez mais. As dificuldades para o acesso à informação científica atualizada nos países subdesenvolvidos, e a pouca visibilidade internacional de sua própria atividade cientifica são duas caras de uma mesma moeda de marginalização e iniqüidade. [...] Promover as potencialidades do Movimento de Acesso Aberto entre a comunidade científica além de trocar seus sistemas de recompensa, contribuirá para transformar o modelo atual de comunicação cientifica em um modelo mais justo e que traga mais benefícios para a sociedade.[4]

É nesse contexto que surgiu a proposta de implantação do Repositório Institucional da UFBA e tendo como comunidade piloto a Editora da universidade por ser a unidade de convergência da produção da instituição, interdisciplinar e aglutinadora que interage com todas as áreas da Universidade.

DESENVOLVIMENTO DO PROJETO

O processo de instalação do DSpace na UFBA, se iniciou em setembro de 2007. Um processo lento, sem uma definição de responsabilidade de um técnico do CPD para esta tarefa e com dúvidas frequentes, solu-

4 Los países subdesarrollados dependen del uso intensivo de los resultados de laactividad científico y técnica para hacer avanzar sus sociedades. Sin embargo, la brecha existente entre estos países y los desarrollados con respecto al acceso, creación y utilización de los conocimientos científicos se ahonda cada vez más. Las dificultades para el acceso a la información científica actualizada en los países subdesarrollados, y la poca visibilidad internacional de su propia actividad científica, son dos caras de uma misma moneda de marginalización e inequidad. [...] Promover las potencialidades del Movimiento Open Access entre la comunidad científica, además de cambiar sus sistemas de recompensa, contribuirán a transformar el modelo actual de comunicación científica en un modelo más justo y beneficioso para la sociedad. (tradução da autora)

cionadas por um técnico do Ibict, Milton Shintaku e a participação de Meirelles. A efetivação deste processo só ocorreu, na prática, em 2008, sendo que em novembro desse mesmo ano foi necessário migrar para a nova versão do DSpace. Nesta etapa, a participação de Meirelles e Shintaku junto ao técnico do CPD foi imprescindível para a finalização da instalação. A partir dessa parceria foi traçado um planejamento para execução do processo de customização que consistiu no seguinte:

- Atualização da versão do Dspace: o sistema foi atualizado para versão 1.5.1 com banco de dados PosGreSql e sistema operacional Linux, instalado em um servidor da UFBA;
- Customização do formulário de entrada: consistiu em adequar o formulário padrão do DSpace às necessidades de descrição dos registros a serem depositados no RI de acordo com as políticas de conteúdo do RI. Nesta etapa foram definidos o levantamento dos metadados, associação dos metadados com o padrão Dublin Core e ajustes nos metadados do formulário padrão do DSpace para entrada dos dados;
- Implantação do tema: foram realizados ajustes para adequar o *layout* do sistema à proposta do RI. Nessa etapa optou-se por não realizar grandes mudanças na estrutura do sistema, pois o mesmo já possui uma arquitetura funcional e a manutenção da estrutura padrão também facilita a inserção de novos recursos e a atualização de versão;
- Configuração das comunidades: foi elaborada uma estrutura baseada nas políticas de conteúdo, de submissão e preservação de acordo com os recursos disponíveis no sistema. Com base nessa estrutura o RI é organizado a partir de comunidades – unidades orgânicas (Escolas, departamentos e centros de pesquisa). Cada comunidade pode organizar os seus documentos em diferentes coleções individualmente, com base na política do RI.

Concluída essa etapa com relação ao Dspace, finalizada a adequação do *layout* e criada a página principal do RI, iniciou-se os testes a partir do que se estabeleceu como projeto piloto com a disponibilização do catálogo de livros da EDUFBA. Para tanto o Conselho Editorial definiu a política de arquivamento para esse conteúdo:

- O Conselho Editorial da Editora da UFBA será soberano nas decisões relativas ao arquivamento dos conteúdos produzidos pela Editora;
- Todo e qualquer conteúdo, já publicado, para ser disponibilizado deverá ter a concordância do autor(es) ou organizador(es), mediante assinatura de um termo. Os autores com conteúdos a serem publicados, no ato de assinatura do contrato de direito autoral, estarão autorizando sua disponibilização, obrigatoriamente.
- Os livros esgotados e relevantes, após análise do Conselho Editorial, serão disponibilizados na íntegra;
- As Coleções MANUAIS e SALA DE AULA além de outras que surgirem tendo como objetivo dar suporte aos cursos de graduação e pós-graduação serão disponibilizadas na integra, independente da data de lançamento;
- Títulos definidos pelo Conselho Editorial, a partir da análise do parecer, que se destina a um público muito restrito, ficarão disponíveis apenas no RI;
- Os novos títulos, excetuando as coleções definidas anteriormente, serão arquivados no RI após seis meses de lançamento, salvo nos casos em que editais e/ou contratos recomendem a obrigatoriedade do acesso livre.

Em 12 de dezembro de 2008, aconteceu no Rio de Janeiro o *I Workshop sobre Acesso Livre nas Unidades de Pesquisa*. O evento teve apoio da Subsecretaria de Coordenação das Unidades de Pesquisa (SCUP) e da Comissão Nacional de Energia Nuclear (CNEN). Na ocasião foi apresentada a Carta do Rio de Janeiro, que teve a assinatura de 10 instituições de ensino e pesquisa e outras instituições presentes se comprometeram em discutir no âmbito de suas instituições e Dr. Hélio Kuramoto, do Ibict fez uma explanação, sobre o Edital para chamadas de projeto para a distribuição de 70 servidores à instituições públicas de ensino e pesquisa. Com este Edital deu-se o primeiro passo para a implantação de ações contemplando o acesso livre à produção científica das unidades de pesquisa. As instituições contempladas através desse e Edital receberiam um servidor já com o DSpace instalado. Logo que publicado o referido Edital decidiu-se que mesmo com o Dspace

já instalado no servidor da Universidade a UFBA participaria, e, caso contemplada teria um servidor exclusivo para o RI com as instalações do DSpace bem como um kit tecnológico completo para atender inclusive a estatísticas relativas ao Repositório. Encaminhou-se a proposta e em março de 2009 a UFBA recebeu o convite do Ibict para ser uma das Instituições piloto do projeto.

O lançamento do edital no mês de dezembro, não alcançou o êxito esperado. Nesse período as instituições de ensino superior do Brasil entram em recesso de final de ano e em seguida ocorrem as férias de verão. Além disso, algumas imperfeições foram detectadas e o Edital Finep/PCAL/SBDB n. 001/2009 foi revogado pois não atendia aos objetivos dos projetos PCAL e XBDB. Uma nova chamada foi publicada e, consequentemente, um novo cronograma foi definido. Portanto, na mesma data, 17/04/2009, foi publicada uma nova chamada de edital, o Edital Finep/PCAL/XBDB n. 002/2009, com prazo para o dia 01 de junho de 2009. Foram contempladas 27 instituições de ensino superior e\ou pesquisa[5] e outras seis foram escolhidas para participarem do projeto piloto[6], dentre eles o Ibict e a UFBA. O kit tecnológico[7] foi entregue aos responsáveis indicados pelas Instituições contempladas e que

5 Centro de Tecnologia da Informação Renato Archer; Fundação Centro Tecnológico de Minas Gerais; Fundação Santo André; Fundação Universidade Federal do Rio Grande; Instituto Nacional de Metrologia, Normalização e Qualidade Industrial; Instituto Nacional de Tecnologia; Museu Paraense Emílio Goeldi; Universidade Tecnológica Federal do Paraná; Universidade de São Paulo; Universidade Estadual de Ponta Grossa; Universidade Federal de Acre; Universidade Federal de Alagoas; Universidade Federal de Goiás; Universidade Federal de Grande Dourados; Universidade Federal de Mato Grosso; Universidade Federal de Ouro Preto; Universidade Federal de Pelotas; Universidade Federal de São Carlos; Universidade Federal de Sergipe; Universidade Federal de Uberlândia; Universidade Federal de Viçosa; Universidade Federal do Maranhão; Universidade Federal do Rio Grande do Norte; Universidade Federal do Tocantins; Universidade Federal dos Vales do Jequitinhonha e Mucuri; Universidade Federal Fluminense; Universidade Municipal de São Caetano do Sul.

6 Instituto Brasileiro de Informação em Ciência e Tecnologia; Universidade de Brasília; Universidade Federal da Bahia; Universidade Federal de Pernambuco; Universidade Federal do Rio Grande do Sul; Universidade Federal de Santa Catarina.

7 O kit tecnológico é composto de um servidor DELL PowerEdge T300, processador Xeon Quad Core, com 4Gb de memória RAM e 2 HDs Satã de 512Gb, RAID 1 (Espelhados). Sistema operacional baseado na Plataforma UNIX\LINUX. Softwares: DSpace 1.5.1, instalado com o embargo e as estatísticas da Universidade do Minho e o SEER e pacotes: Apache 2.1 (um), PHP 5, MySQL, Postgresql, Tomcat 6, Java versão JDK 1.5, Pacotes Java-Ant, Ant-optional e Maven 2.

participaram do *I Encontro sobre Gestão de Repositórios Institucionais* no dia 6 de novembro de 2009 e o cronograma estabelecido foi o seguinte:

CRONOGRAMA									
ATIVIDADES	MESES								
	2009		2010						
	11	12	01	02	03	04	05	06	07
01 - Planejamento									
02 - Desenvolvimento e implantação									
03 - Discussão e estabelecimento da política									
04 - Articulação e operalização									
05 - Lançamento e operacionalização									

Fonte: Apresentado pelo Ibict no I Workshop sobre Acesso Livre nas Unidades de Pesquisa.

Para a implantação de uma política institucional para respaldar esse projeto na UFBA, foi criado um grupo gestor designado através da Portaria n. 189, de 1 de abril de 2009, com representantes do Instituto de Ciência da Informação (ICI), EDUFBA, Faculdade de Comunicação (FACOM) e CPD e, em janeiro de 2010, passa a fazer parte também um representante do Sistema de bibliotecas. Este grupo foi encarregado de discutir e propor a política de depósitos para o RI e as ações instituicionais para esta implantação. Entende-se que a gestão passa não apenas pelos critérios, bem como definição de aporte financeiro, definição de linhas de trabalho para sensibilização da comunidade científica da Universidade. No dia 30 de setembro foi encaminhada a proposta para análise e assinatura da política para o RI da UFBA e, 7 de janeiro de 2010, o Reitor, Prof. Dr. Naomar de Almeida Filho, assinou a Portaria n. 24/2010 que estabelece esta política de arquivamento.

CONCLUSÃO

Diferentemente da Universidade do Minho que tem servido de inspiração e orientação para o projeto da UFBA, não houve uma participação institucional no que diz respeito ao movimento mundial do *Open Access.* O Budapest Open Access Initiative (BOAI) foi divulgado em 2001 e em 2003 tinham início as atividades para a implantação do RepositóriUM, demonstrando a articulação dos responsáveis pelo projeto com as novas tendências internacionais em termos de acesso livre. Desse modo, acredita-se que há necessidade de sensibilizar e preparar a comunidade da UFBA para essa nova realidade mundial.

Considera-se também que, para que a Universidade Federal da Bahia tenha o sucesso almejado no seu projeto, é evidente a necessidade de um compromisso da administração central da Instituição, tanto no que diz respeito à política como também aos recursos. A mobilização em torno da questão é importante para que os pesquisadores tenham consciência da necessidade, hoje, de adesão ao *Open Access,* acompanhando as mudanças na forma de disseminação da produção científica.

Uma questão tecnológica importante diz respeito ao sistema de estatística[8] - Add-on - desenvolvimento pela Universidade do Minho que apresenta funcionalidades tecnológicas compatíveis com o Dspace e permite o registo e disponibilização de dados estatísticos. Graças a este programa é possivel se obter dados referentes ao número de consultas e *downloads* de documentos, a origem geográfica desse processo de consulta/*download*, ter uma noção do impacto do Repositório tanto na visibilidade da instituição e da sua produção e publicação científicas, como também permite que o depositante acompanhe o comportamento da sua produção que foi disponibilizada no RI.

A inclusão da UFBA no projeto piloto do Ibict referente ao Edital Finep/PCAL/SBDB n. 001/2009, substituído pelo Edital Finep/PCAL/XBDB n. 002/2009 criou um compromisso maior da Instituição com relação a implantação do RI da UFBA, quer seja com relação ao cronograma que passou a ser adotado, quer seja pelo compromisso com uma política

8 StatisticsAddOn, disponível em: http://wiki.dspace.org/index.php/StatisticsAddOn

que de fato traga os resultados pensados para uma política nacional de disseminação da produção científica brasileira.

Finalmente, é indispensável o investimento na promoção e divulgação do RI, seja através de apresentações internas na própria instituição, seja em outros contextos e países. No caso do RepositóriUM, uma iniciativa importante é a sua integração no projeto do governo português, Repositório Científico de Acesso Aberto de Portugal (RCAAP), que tem como objetivo "aumentar a visibilidade, acessibilidade e difusão dos resultados da actividade académica e de investigação científica nacional e facilitar o acesso à informação sobre a produção científica nacional em regime de 'open access', bem como integrar Portugal num conjunto de iniciativas internacionais neste domínio". (RCAAP, 2009). No caso do Brasil, a referência é o Ibict que tem desenvolvido políticas de estímulo ao acesso livre ao conhecimento disponibilizando e dando subsídios tecnológicos para o uso de *softwares* livres voltados para a disseminação da produção científica como o Sistema Eletrônico de Editoração de Revistas (SEER), Sistemas de Arquivos Digitais (D-SPACE e DiCi) além da criação do Portal Brasileiro de Repositórios e Periódicos de Acesso Livre (OASIS.Br).

REFERÊNCIAS

A AVENTURA EDITORIAL DE PINTO DE AGUIAR. Salvador: Instituto Baiano do Livro, 1993.

DIAS, André Luis Mattedi. A universidade e a modernização conservadora na Bahia: Edgard Santos, o Instituto de Matemática e Física e a Petrobras. **Revista SBHC**. Rio de Janeiro, v. 3, n. 2, p. 125 – 145. jul.|dez. 2005. Disponível em: <http://www.mast.br/arquivos_sbhc/7.pdf> Acesso em: 27 jan. 2010.

FERRAZ, Marcelo Carvalho. Nota prévia. In: RISÉRIO, Antonio. **Avant-garde na Bahia**. São Paulo: Instituto Lina Bo Bardi, 1995.

GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. São Paulo: Atlas, 1999.

LYNCH, Clifford A. Institutional Repositories: essential infrastructure for scholarship in the Digital Age. *ARL Bimonthly Report*. 2003 p. 1-7. Disponível em: < http://www.arl.org/newsltr/226/ir.html>. Acesso em: 20 ago. 2008.

NASCIMENTO, Antonio Dias et al. **Subsídios para uma política de comunicação da UFBA**. Salvador, 1993, 32 p. Fotocópia.

RCAAP - Repositório científico de acesso aberto de Portugal. 2009. Disponível em: <http://projecto.rcaap.pt/>. Acesso em: 23 out. 2003

RODRIGUES, Eloy et al. RepositóriUM : criação e desenvolvimento do Repositório Institucional da Universidade do Minho. In: CONGRESSO NACIONAL DE BIBLIOTECÁRIOS, ARQUIVISTAS E DOCUMENTALISTAS, 8., Estoril, 2004. **Actas**... Lisboa: Associação Portuguesa de Bibliotecários, Arquivistas e Documentalistas, 2004. [CD-ROM].

ROSA, Flávia Goulart Motta. Garcia. **A editoração na Universidade Federal da Bahia**. Salvador, UCSAL/FUNDESP, 1994.

RUBIM, Antonio Albino Canelas (Coord.). **Universidade da Bahia** – a ousadia da criação. Salvador: Universidade Federal da Bahia, Faculdade de Comunicação, 1999.

SANCHEZ TARRAGÓ, N. La comunicación de la ciência em los países em vias de desarrollo y el Movimiento Open Access. **Biblios**, ano 8, n. 27, ene./mar., 2007.

SPINK, Peter Kevin. Pesquisa-ação e análise de problemas sociais e organizacionais complexos. **Revista de Psicologia**, Fortaleza, v. 5, n. 1, p. 31-34, 1976.

THIOLLENT, Michel. **Metodologia da pesquisa-ação**. 14. ed. aum. São Paulo: Cortez, 2005.

VIEIRA, Sonia Chaga. **Revistas científicas**: estudo da visibilidade das revistas publicadas pela Universidade Federal da Bahia – UFBA. 2006. 180f. Dissertação (Mestrado em Ciência da Informação) – Instituto de Ciência da Informação, Universidade Federal da Bahia, Salvador, 2006.

UFBA - UNIVERSIDADE FEDERAL DA BAHIA. Departamento Cultural da Reitoria. **Documentos históricos.** Salvador: UFBA, 1971.

UFBA - UNIVERSIDADE FEDERAL DA BAHIA. Departamento Cultural da Reitoria. **Notícia histórica da Universidade da Bahia**. Salvador, 1967.

______. Biblioteca Digital UFBA. **Portaria n. 332, de 8 de julho de 2002. Estabelece o depósito obrigatório de toda a produção científica da Universidade Federal da Bahia na Biblioteca Central Reitor Macedo Costa / Seção Memória da UFBA. Salvador, 2002.** Disponível em: <http://www.bdtd.ufba.br /Portaria%20332.pdf>. Acesso em: 17 ago. 2009.

______. **Portaria n. 80, de 8 de março de 2005. Cria a Biblioteca Digital da Universidade Ferderal da Bahia. Salvador, 2005a. Disponível em: <**http://www.bdtd.ufba.br/ portaria080_05.pdf**>. Acesso em: 10 set. 2009.**

LEXML:
Portal especializado em informação jurídica e legislativa

João Alberto de Oliveira Lima

INTRODUÇÃO

A Constituição Federal de 1988, no art. 5º, inciso XIV, define o acesso à informação como um direito fundamental do cidadão. Vários órgãos disponibilizam informações legislativas e jurídicas em sítios da Internet. A criação de uma infraestrutura que permita organizar e integrar a grande quantidade de informações legislativas e jurídicas existente facilita o acesso à informação e contribui para a realização desse direito fundamental. Nesse sentido, foi criado o *Projeto LexML Brasil* que é uma iniciativa conjunta de diversos órgãos participantes da Comunidade TI Controle em busca do estabelecimento de padrões abertos, integração de processos de trabalho e compartilhamento de dados de interesse comum, que permitam a identificação e a estruturação de informações legislativas e jurídicas.

Um primeiro resultado do *Projeto LexML* é o Portal da Rede de Informação Legislativa e Jurídica[1], inaugurado em 30 de junho de 2009 com 1,28 milhão de documentos[2]. Trata-se de um ponto de acesso unificado às diversas fontes de informação legislativa e jurídica do governo, criando assim a infraestrutura necessária para a difusão em larga escala do identificador unívoco e persistente. Além do portal, estão sendo desenvolvidas ferramentas, tais como, editor, compilador e consolidador com o objetivo de auxiliar a gestão da informação documental.

---

1 http://www.lexml.gov.br

2 Em março de 2010, o acervo era formado por 1,4 milhão de documentos.

Na organização de um acervo é necessário definir, entre outras coisas, um critério de identidade para as classes de interesse, ou seja, para cada tipo de item devemos escolher as características que identificam univocamente cada instância. Por exemplo, no caso de normas jurídicas, é possível considerar os seguintes tipos de itens: a) a norma de uma forma geral, considerando toda a sua evolução no tempo; b) uma versão do texto da norma para uma determinada data; c) um dispositivo de uma versão específica da norma. Para cada um destes três níveis de granularidade listados, é possível definir um grupo de elementos (datas, tipos, números sequenciais etc.) para a composição de um identificador unívoco. Raciocínio análogo aplica-se às proposições legislativas e aos julgados. Uma das principais vantagens de se definir claramente identificadores unívocos é a possibilidade de se estabelecer remissões textuais precisas.

As recomendações do *Projeto LexML* reúnem as melhores práticas de iniciativas para identificação e estruturação de documentos legislativos e jurídicos, adaptando as experiências bem-sucedidas do exterior às necessidades do nosso ordenamento jurídico e da nossa técnica legislativa. Devido à influência do Direito Romano na organização do nosso ordenamento jurídico, as iniciativas de países que adotam o sistema *civil law* são mais aplicáveis à nossa realidade do que as de países que adotam outros sistemas legais. A Itália é um dos países com maior evolução na aplicação de novas tecnologias para a estruturação e organização de normas jurídicas no formato XML. O Projeto *Norme in Rete* (ARCHI et al. 2000), coordenado pelo Centro Nazionale per Informatica nella Pubblica Amministrazione (CNIPA), e realizado com o suporte do Istituto di Teoria e Tecniche dell'Informazione Giuridica (ITTIG) e da Universidade de Bolonha, é um projeto implantado desde 2002 com resultados expressivos. O *Akoma Ntoso* (VITALI, 2010) é um projeto promovido pelas Nações Unidas e realizado com o suporte de pesquisadores da Universidade de Bolonha com o objetivo de capacitar os parlamentos da África com novas tecnologias de informação e comunicação.

Esse capítulo esta organizado em três seções. A primeira apresenta a interface do Portal LexML; em seguida, são apresentados alguns

aspectos da coleta de metadados; e, por fim, é apresentado o identificador unívoco e persistente, fundamental para a interoperabilidade semântica.

PORTAL LEXML – REDE DE INFORMAÇÃO LEGISLATIVA E JURÍDICA

A página inicial do Portal LexML é simples e apresenta um campo para digitação do argumento de pesquisa. É possível, nessa tela, restringir a pesquisa à legislação, à jurisprudência ou às proposições legislativas. A figura 1 apresenta a página inicial onde foi digitado o termo "Lei Maria da Penha" para pesquisa. A figura 2 apresenta o resultado da pesquisa formulada com seis documentos.

Figura 1 - Página Inicial do Portal LexML

Figura 2 - Resultado de pesquisa do termo "Lei Maria da Penha"

O resultado da pesquisa vem acompanhado de filtros que permitem o refinamento da pesquisa de acordo com as suas necessidades de informação. Cinco filtros estão disponíveis: localidade, autoridade emitente, tipo de documento, data e siglas.

Figura 3 - Detalhe do primeiro documento do resultado de pesquisa

A página de detalhe do documento (Figura 3) apresenta os *links* da publicação oficial (Imprensa Nacional) e os *links* de outras publicações disponíveis dos sistemas de informações da Presidência da República, da Câmara dos Deputados e do Senado Federal. Também são apresentados *links* com o nome "Mais detalhes" que remetem para a página de metadados da Câmara dos Deputados e do Senado Federal e *links* com o nome "Linker" que ativam o serviço de identificação automática de remissões textuais.

COLETA DE METADADOS

A coleta de metadados tem por objetivo reunir os metadados de documentos legislativos e jurídicos disponíveis nos sítios dos diversos

órgãos governamentais. São coletados, inicialmente, metadados de identificação (epígrafe, apelidos, identificadores etc.) e metadados descritos (ementa e relacionamentos). Como forma de facilitar e automatizar o processo de coleta de metadados foi escolhido o Protocolo OAI-PMH (Open Archives Information - Protocol for Metadata Harvest).

A arquitetura de uma rede de informações que utiliza o Protocolo OAI-PMH para intercâmbio de metadados é formada por nodos de três tipos (Figura 4):

- Provedor de Dados (*Data Provider*) – serviço responsável pela exposição de metadados;
- Provedor de Serviço (*Service Provider*) – serviço responsável pela comunicação com os nodos provedores e agregadores de dados, pelo processamento dos dados coletados e pela oferta de serviços de pesquisa;
- Agregador de Dados (*Data Aggregator*) – serviço responsável por agregar metadados coletados de provedores de dados e disponibilizá-los para um provedor de serviço. No LexML, o Tribunal Superior do Trabalho e o Tribunal Superior Eleitoral foram definidos como agregadores das informações da justiça trabalhista e eleitoral, respectivamente.

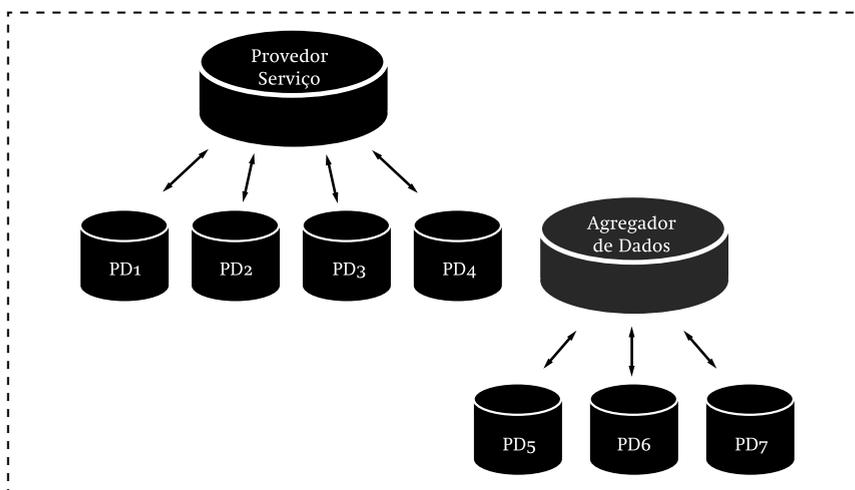



Figura 4 - Arquitetura OAI-PMH

Para cada provedor de dados é possível definir no LexML o papel de “Publicador”. Por exemplo, inicialmente no Senado Federal como provedor de dados foram definidos três publicadores: um para as normas jurídicas; um outro para proposições legislativas e um terceiro para os diários oficiais do Senado Federal e do Congresso Nacional.

O protocolo OAI-PMH caracteriza-se pela simplicidade dos comandos (apenas 6 verbos) e pela fácil integração a qualquer ambiente computacional, pois é baseado apenas no Hypertext Transfer Protocol (HTTP) e no formato eXtensible Markup Language (XML).

Cada registro de metadado é composto por um cabeçalho (dados de identificação do protocolo), um corpo (metadado propriamente dito) e, opcionalmente, uma seção com informações de proveniência do registro.

A figura 5 apresenta um exemplo de um registro de metadado do LexML que é coletado pela arquitetura proposta.

```
<LexML
    xsi:schemaLocation=”http://projeto.lexml.gov.br/esquemas/
oai_lexml.xsd”
    xmlns =”http://www.lexml.gov.br/oai_lexml”
    xmlns:xsi=”http://www.w3.org/2001/XMLSchema-instance”>

    <Item formato=”text/html”>
       http://www6.senado.gov.br/legislacao/ListaPublicacoes.
action?id=102415
    </Item>
    <DocumentoIndividual>
     urn:lex:br:federal:lei:1990-09-11;8078@1990-09-12!1990-09-
12~texto;pt-br
     </DocumentoIndividual>
     <Epigrafe>Lei nº 8.078, de 11 de setembro de 1990</
Epigrafe>
     <Apelido>Código de Defesa do Consumidor</Apelido>
     <Apelido>Código de Proteção e Defesa do Consumidor</
Apelido>
     <Apelido xml:lang=”es”>
        Código de Protección y Defensa del Consumidor
     </Apelido>
     <Ementa>
        Dispõe sobre a proteção do consumidor e dá outras
providências
     </Ementa>
    <Relacionamento tipo=”publicacao.oficial”>
urn:lex:br:imprensa.nacional:publicacao.oficial;diario.oficial.
uniao;secao.1:1990-09-12;123:pag1
    </Relacionamento>
</LexML>
```

Figura 5 - Exemplo de Registro de Metadados OAI-LEXML

O elemento `<Item>` possui a URL do recurso disponibilizado na Internet sob a responsabilidade do publicador. A cada `<Item>` deve-se, obrigatoriamente, relacionar uma URN no elemento `<DocumentoIndividual>`. Esse relacionamento posiciona o recurso publicado no espaço de nomes definidos pela Parte 2 das especificações LexML. Na sequência, são relacionados: a epígrafe, os apelidos registrados para o documento e a ementa. Por fim, sempre que possível, deve-se relacionar ao documento individual identificado a URN da publicação oficial que o veiculou.

Os elementos textuais (epígrafe, apelido e ementa) possuem o atributo `xml:lang` para especificação da língua. Esse atributo tem por default o valor "pt-BR".

Caso uma URL contenha mais de um documento individual (por exemplo, a norma e seus anexos), deve-se criar registros adicionais para cada documento individual relacionado.

O elemento `<Relacionamento>` permite estabelecer associações entre documentos legislativos e jurídicos, especificado pelo atributo "`tipo`".

IDENTIFICADOR UNIFORME – URN

A remissão textual é muito utilizada nos documentos legislativos e jurídicos. Por exemplo, uma norma jurídica pode fazer referências a outras normas, seja para realizar alterações, revogar um dispositivo ou regulamentar uma matéria; um acórdão pode referenciar normas jurídicas bem como outros acórdãos e súmulas.

As remissões textuais podem ser classificadas de diversas formas:

- Quanto à localização do alvo de uma remissão, podemos identificar dois tipos: as remissões internas e as externas. As remissões internas têm como alvo o próprio documento. As externas referenciam outros documentos;
- Quanto à granularidade de uma remissão externa, podemos identificar dois tipos: a remissão ao documento no todo e a remissão a uma parte do documento;

- Quanto à dimensão temporal de uma remissão externa, podemos identificar dois tipos: a remissão a uma versão específica de uma norma e a remissão à norma de uma forma genérica, sem especificar uma versão específica.

Para representar corretamente as remissões é necessário criar um identificador que qualifique unicamente cada possível alvo de uma remissão textual. Na impossibilidade de criar um sistema único de numeração para todo o acervo de normas, julgados e proposições legislativas do Brasil, fez-se necessária a criação de um identificador que seja compatível com os diversos sistemas de numeração existentes.

O Projeto LexML define um identificador unívoco e persistente para os documentos legislativos e jurídicos. A persistência possibilita que os alvos sejam encontrados, mesmo que ocorra mudança de endereços da Internet – *Universal Resource Locator* (URLs) –, evitando assim o conhecido "Erro 404" (recurso não encontrado) do protocolo HTTP. Para que a persistência se realize é necessário que apenas os provedores de dados atualizem a nova localização do recurso quando da mudança de localização.

O identificador LexML permite a integração entre as fontes de informações legislativas e jurídicas, pois define um referencial único para esse tipo de documento.

A lista abaixo relaciona endereços da Internet (URLs) onde é possível consultar o texto da Lei n. 8.666/1993 nos sítios do Senado Federal e da Presidência da República, respectivamente. É possível notar que cada identificador possui regras de formação diferentes e está diretamente relacionado à solução tecnológica adotada pela instituição em um determinado momento.

http://www6.senado.gov.br/legislacao/ListaPublicacoes.action?id=102446

http://www.planalto.gov.br/ccivil/Leis/L8666cons.htm

Caso ocorra mudança na tecnologia utilizada ou alteração na convenção de nomes por parte dos órgãos gestores desses recursos, todos os *links* dos documentos que referenciavam esses identificadores irão apresentar Erro 404 (recurso não encontrado) ao serem ativados.

O identificador LexML utiliza a notação URN que permite a criação de identificadores persistentes, pois estes não estão atrelados a detalhes de tecnologia ou à uma convenção de nomes definida de forma unilateral. As URNs seguem um esquema previamente acordado entre os participantes da rede. Abaixo, é apresentada a URN da Lei n. 8.666/1993.

urn:lex:br:federal:lei:1993-06-21;8666

PRINCÍPIOS DO NOME UNIFORME

O nome uniforme deve ser unívoco, isto é, deve identificar uma e apenas uma entidade, e é construído, tanto quanto possível, alinhado com os princípios apresentados a seguir:

- Princípios Gerais da URN
  - autoexplicativo para os usuários;
  - dedutível por meio de regras simples e claras;
  - alinhado constantemente aos demais padrões do projeto.
- Princípios da URN Canônica
  - unívoco para o espaço das URNs canônicas;
  - registrável com o mínimo de informação (o suficiente para ser unívoca);
  - representativo dos aspectos formais do documento;
  - em conformidade, à data de sua emissão, com a estrutura/organização da autoridade emitente e com a tipologia do documento.
- Princípios da URN de Referência
  - representativo dos aspectos formais e/ou substanciais do documento;
  - mapeável em URNs canônicas cabíveis;
  - compatível com a prática em uso para criar referências;
  - reduzido ao essencial, para simplificar os *links* com outros documentos;
  - capaz de ser gerado de forma automática por analisadores de remissões textuais;

ESTRUTURA BÁSICA DE URNS E EXEMPLOS

O nome uniforme do *Projeto LexML Brasil* é prefixado por “urn:lex:” e é seguido, basicamente por mais quatro elementos, delimitados por “:” (símbolo de dois pontos), conforme lista abaixo:

- localidade – contém a jurisdição da autoridade emitente;
- autoridade – indica a autoridade emitente do documento;
- tipo do documento – indica a espécie da tipologia documental;
- descritor – formado pela data representativa e um descritor alfanumérico.

Os três primeiros elementos são definidos em vocabulários controlados que estão sendo construídos à medida em que as instituições aderem ao *Projeto LexML*.

A lista abaixo relaciona alguns exemplos de nomes uniformes de documentos jurídicos e legislativos que estão disponíveis no Portal LexML.

- Lei n. 11.705, de 19 de Junho de 2008.
    - urn:lex:br:federal:lei:2008-06-19;11705
- Súmula Vinculante n. 11
    - urn:lex:br:supremo.tribunal.federal:sumula.vinculante:2008-08-13;11
- Lei nº 13.745 de 10 de fevereiro de 2006 (Lei Municipal – São Carlos – SP)
    - urn:lex:br;sao.paulo;sao.carlos:municipal:lei:2006-02-10;13745
- Acórdão TCU 395/1997
    - urn:lex:br:tribunal.contas.uniao;camara.1:acordao:1997-09-02;395

TIPOS DE NOMES UNIFORMES

Para alcançar uma maior precisão nas gramáticas de definição das URNs, foram definidos dois tipos de URN:

- URN Canônica – É a URN normalizada segundo o vocabulário e

regras mais restritas, existindo apenas uma para cada entidade identificável (documento, documento complexo e fragmento). Ela é obrigatoriamente informada pelo provedor de informações e relacionada a um documento específico.

- URN de Referência – É a URN utilizada pelos usuários ou sistemas de informação nas referências a documentos, podendo utilizar a forma compacta, omitir elementos que assumirão valores *default* ou ainda utilizar nomes alternativos.

CONCLUSÃO

A informação legislativa e jurídica é altamente caracterizada por relacionamentos. A regra geral é que qualquer documento deste domínio está sempre inserido em um contexto, parte do nosso ordenamento jurídico. O *Projeto LexML* adapta para o Brasil experiências bem-sucedidas do exterior que inclui a definição de um identificador persistente implementado como uma URN e a criação de um Portal unificado de acesso. O identificador é utilizado como instrumento para representar as referências (e, mais genericamente, quaisquer tipos de relações) entre os documentos eletrônicos, com a finalidade de tornar disponível um ambiente hipertextual global sobre os recursos de informação, operacional em um ambiente distribuído. Mais detalhes sobre o *Projeto LexML* e a definição das URN, incluindo a gramática em formato Extended Backus Naur Form (EBNF), podem ser consultados no Portal de Documentação do Projeto.[3]

Os esforços de criação do Portal e de definição do identificador unívoco e persistente não são suficientes para se atingir o objetivo principal: organização da informação legislativa e jurídica disponível em sítios governamentais na Internet do Brasil. De forma complementar, o *Projeto LexML* está desenvolvendo ferramentas (*software* livre) tais como editores, compiladores, consolidadores, comparadores, entre outras para auxiliar a gestão da informação documental, catalisando,

3 http://projeto.lexml.gov.br

dessa forma, a capacitação informacional e tecnológica dos órgãos da Rede LexML.

REFERÊNCIAS

ARCHI, A. et al. Studio di fattibilità per la realizzazione del progetto “Accesso alle norme in rete”. **Informatica e Diritto**. v. 24, n. 1. P. 1-200, 2000.

VITALI, F. **Akoma Ntoso Release Notes**. 2008. Disponível em: <http://www.akomantoso.org>. Acesso em: 5 mar. 2010.

RI é a resposta, mas qual é a pergunta?

Primeiras anotações para a implementação de Repositório Institucional

Maria Cristina Soares Guimarães
Cícera Henrique da Silva
Ilma Horsth Noronha

INTRODUÇÃO

No começo dos anos 1980, um artigo publicado por um influente pesquisador britânico da área de comunicação no *Journal of Information Science* (HALLORAN, 1983) teve como título a seguinte provocação: "Information may be the answer, but what is the question?". Testemunha de tempos quando se forjou a expectativa que as tecnologias de informação e comunicação (TICs) seriam um instrumento poderoso e uma estratégia irresistível a favor da democracia e da ampliação da participação social nos processos de decisão do Estado, o autor se perguntava se não havia ali um otimismo excessivo. Ou, disponibilidade de informação é o que realmente importa? Ainda, disponibilidade de informação é condição necessária à democracia, mas seria suficiente? Os questionamentos do autor ecoam, ainda, aquelas perguntas clássicas, tão caras aos profissionais de informação, especialmente àqueles dedicados ao desenvolvimento de Sistemas de Recuperação de Informação (SRI): quem necessita de informação? Que tipo de informação diferentes grupos de usuários, ou toda a população, necessita? Quem decide o que é necessário para quem? Quem seleciona, organiza e provê acesso à informação? A que custo? Que uso é esperado ser feito da informação?

Passadas quase três décadas, período esse de avanços extraordiná-

rios no âmbito das TICs, permanece (ou mesmo fortalece) a aposta que a mais ampla circulação de informação tem um papel fundamental no bem estar social das nações, especialmente dado o caráter central e produtivo do conhecimento científico e tecnológico. Acesso à informação tornou-se uma força propulsora do processo de inovação. A Internet e a descentralização do processo de produção e disseminação do conhecimento, junto a uma crescente prática de colaboração em rede são forças modeladoras da atual reconfiguração da prática da ciência, e de seus processos de comunicação e publicação. Nasce aqui uma oportunidade única para uma resposta construtiva às disfunções do mercado que operam no sistema de comunicação científica – o conhecimento como bem público é uma utopia possível (GUIMARÃES, 2009).

O uso intensivo das mídias eletrônicas tornam virtualmente impossível a limitação do acesso à informação, sob pena do uso de soluções jurídicas antipáticas e draconianas, que só fazem aprofundar o gap entre os have e os have not. A resistência veio pelo Movimento do Acesso Livre à informaçao científica: "O propósito da informação é para ser compartilhada, assim como o propósito do pão é para ser comido"[1](OPEN SOCIETY INSTITUTE, 2005, tradução nossa) .

Repositório Institucional (RI), sumariamente entendido como um locus online para coleta, preservação e disseminação da produção intelectual de uma instituição (de pesquisa), emerge, assim, como uma peça fundamental no quebra-cabeça da iniquidade no acesso à informação científica. O que começou no campo disciplinar, no seio da comunidade científica, como uma estratégia para intercâmbio de preprints, nos anos noventa do século passado, chega ao século XXI com ares de inevitabilidade: "[...] É improvável que qualquer instituição séria de pesquisa não vá ter seu próprio repositório digital até ao final desta década."[2] (SWAN; CARR, 2008, p. 31).

Reverberando a provocação de Halloran (1983), e guardadas as devidas proporções, as discussões em torno do conceito, desenvolvimento

1 "The whole purpose of information is to be shared, as the purpose of bread is to be eaten."

2 "It is unlike that any serious research-based insitution will not have its own digital repository by the end of this decade."

e implantação de RIs parecem seguir a mesma dinâmica – dispor de um RI é o que importa! Especialmente na última década, e particularmente no esteio do conjunto de declarações e manifestos, nacional e internacionais, a favor do Acesso Livre (KURAMOTO, 2006), os RIs emergem como resposta para um amplo conjunto de iniquidades no acesso à informação, nas suas mais diversas manifestações. Quer seja como resposta à escalada nos custos das assinaturas de periódicos científicos; como um novo "serviço de informação"; como um elemento adicional no sistema de comunicação cientifica; como uma estratégia para conferir visibilidade institucional, ou mesmo para a preservação digital da memória institucional, os RIs podem, potencialmente, ser a resposta! A despeito de tamanha onda de otimismo, o que a literatura na área aponta é que as experiências em curso testemunham que muito ainda há que ser investigado e colocado em prática para que todo o potencial dos RIs possa ser alcançado (SHEEREVES; CRAGIN, 2008).

Nesse sentido, o que talvez seja ainda pouco explorado, ainda que não esquecido, nas investigações em curso é que muito antes de ser mais um dispositivo ou infraestrutura tecnológica a favor da ampliação da circulação de informação cientifica, um RI é um complexo sociotécnico, onde estão envolvidos um conjunto de atores com visões, práticas e perspectivas diversas, cujas relações são tecidas à luz de micro e macropolíticas, locais e globais. Ainda que a importância dos aspectos técnicos seja inegável para sua concepção, desenvolvimento e implantação, eles não asseguram, ou sustentam, a apropriação e uso do RI pelos usuários reais e potenciais. Como insistentemente lembrado pelas análises empreendidas na área dos estudos sociais das tecnologias, mais particularmente, da construção social das tecnologias (PINCH; BIJKER, 1984): é somente pelo uso, e pela perspectiva do usuário, que uma tecnologia é absorvida no social e reconhecida em sua qualidade e seus méritos.

O texto que se segue objetiva ser uma breve reflexão sobre alguns pontos norteadores para o desenvolvimento e implantação de um RI. No limite entre a inexorabilidade de implantá-lo e a desqualificação de seu potencial no esforço de ampliação da circulação da informação científica, repousa o reconhecimento que um RI é muito mais que

uma contribuição ímpar das TICs em favor do Movimento do Livre Acesso. O "I", de institucional, que adjetiva o repositório (esse, sim, um objeto digital) coloca em foco uma instituição que é um complexo de relações sociais e de práticas singulares, embebidas na cultura e nas contingências políticas e de ordens outras, resultado de sua própria história, e cujas metas e objetivos mais amplos orientam um futuro almejado. Some-se a essa instituição um segundo adjetivo, "de pesquisa", onde domínios de conhecimento, tradições de pesquisa, culturas epistêmicas, práticas de comunicação e padrões de busca (*information seeking*) e uso de informação implicam (e estão implicadas) em relações idiossincráticas com o desenvolvimento, estabilização e uso de aparatos tecnológicos, e com a própria dinâmica da ciência.

Nesse sentido, qualquer visão simplificadora do relacionamento entre tecnologias, espaços de práticas constituídos e mudanças sociais é, no mínimo, ingenuidade. Perspectivas de análise construtivista e sociotécnica, especialmente no âmbito da "construção social das tecnologias" (PINCH; BIJKER, 1984) e da "informática social" (KLING, 2000; 2007), ressaltam a importância de um olhar cuidadoso por sobre os contextos institucional e cultural quando do desenvolvimento e implantação de um RI. Entram em campo, portanto, considerações políticas e todo um repertório de estratégias que devem ser levados em consideração *vis a vis* o problema a ser enfrentado.

Longe de ser um *tour de force* sobre o tema (esforço que, por certo, precisa ser empreendido), a discussão que se segue visa, antes, uma leitura transversal sobre alguns pontos discutidos na literatura sobre a natureza e função dos RIs, bem como suas estratégias de desenvolvimento e implantação, e como a perspectiva sociotécnica pode enriquecer o aprendizado esperado. Aqui, convivem esperanças e incertezas, como as duas faces de Juno: por um lado, lê-se o esforço e comprometimento genuíno de muitos (órgãos governamentais, organizações de classe, instituições de pesquisa, e pesquisadores individuais) para a promoção e implantação de RIs. Por outro lado, permanece a dúvida de como realizar a potência do acesso universal quando práticas estabelecidas, culturas, e interesses econômicos herdados de séculos ainda perduram no cerne da dinâmica cientifica.

Muito antes de procurar por respostas, o texto procura colocar algumas perguntas que, quiçá, possam estimular análises mais detalhadas sobre o tema, especialmente no contexto brasileiro.

A LENTE DA CONSTRUÇÃO SOCIAL DAS TECNOLOGIAS

O campo de investigação "construção social das tecnologias" (Social Construction of Technology – SCOT) (PINCH; BIJKER, 1984) faz parte de um conjunto de abordagens teórico-metodológicas que se aglutinam sob o guarda-chuva do movimento de Social Shaping of Technology – SST (WILLIAM; EDGE, 1996), que advoga que uma tecnologia não é nem puramente tecnológica tampouco emana diretamente de uma estrutura ou demanda social. Ao contrário, ela é concebida, desenvolvida e alcança sua estabilização tecida em uma rede complexa e mestiça de atores, interesses, estratégias e atividades. O ponto central é entender que escolhas são feitas, e que qualquer perspectiva de imperativos ou determinismos (tecnológico, científico, político, econômico ou social) é impotente para descrever como as tecnologias chegaram ao que são em sua forma final. Trata-se, afinal, não de uma tecnologia, mas um dispositivo sociotécnico (GUIMARÃES, 1997).

De comum, essas perspectivas compartilham que:

- O processo de desenvolvimento das tecnologias é sempre contigente;
- As tecnologias nascem sob o conflito, a diferença e a resistência;
- As tecnologias estão implicadas nas estratégias dos atores, e só quando a concordância é possível é que as tecnologias ganham forma e sentido social mais amplo;
- Tanto as estratégias quanto suas consequências são um fenômeno emergente no sistema de relações entre os atores. Ou seja, elas nascem no curso da ação e, recursivamente, se influenciam.

Ao longo dos anos, a perspectiva teórica SCOT mostrou-se bastante produtiva para analisar o desenvolvimento e apropriação das TICs em

diferentes contextos sociais (WILLIAMS, 1997; WILSON; HOWCROFT, 2002), e foi tomada como base para o desenvolvimento do campo de investigação da "informática social", capitaneada por Kling (2000, 2007).

De forma sumária, e para o interesse da presente discussão, a abordagem SCOT defende que o processo de inovação que leva à apropriação e ao uso das tecnologias é complexo, coletivo, e submetido a intensas negociações de sentido, moldadas e informadas pelo contexto no qual os atores se posicionam. Ao longo do processo, mudam tanto o contexto e as relações entre os atores, quanto os artefatos tecnológicos. Quatro conceitos básicos orientam as análises na abordagem SCOT:

- Grupos sociais relevantes, ou, os diferentes atores (*stakeholders*) que possuem lugar de destaque no processo, e para ele trazem diferentes recursos (conhecimento, poder, financiamento, dentre outros), diferentes visões e expectativas sob a forma e sentido futuro dos artefatos;
- Flexibilidade interpretativa, ou, em função da participação de diferentes grupos sociais relevantes, o processo está aberto a várias interpretações de sentido e de uso;
- Quadro tecnológico, ou, um conjunto de conceitos, práticas e técnicas compartilhadas e utilizadas pelos grupos que informam os interesses comuns dos grupos;
- Fechamento ou estabilização, ou, quando o processo de negociação é finalizado pelo consenso e emerge dali um "modelo dominante".

Sob a ótica SCOT, um RI deve primeiramente ser situado em um contexto de mudança e aberto à inovação, como é o caso do setor de publicação acadêmica face às TICs e ao Movimento do Livre Acesso (BORGMAN, 2007; KING et al., 2006; BOHLIN, 2004). Essa configuração delineia o quadro tecnológico. Aqui, são potenciais grupos sociais relevantes, por exemplo, os pesquisadores, as instituições de pesquisa, os editores científicos (comerciais e universitários), as bibliotecas e os profissionais de informação, as agências de fomento, os usuários, o Legislativo e outras organizações, governamentais ou não. Ressalte-se que cada um deles traz para a discussão dos RIs

um conjunto de "recursos" e interesses, que também podem estar longe de representar um consenso interno no grupo. O que cada um expressa como uma macropolítica pode ser, de fato, um conjunto de micropolíticas, por vezes conflitantes. O que, teoricamente, os une em um "quadro tecnológico" é o interesse na ampliação do acesso à informação, ainda que para alguns isso signifique geração de riqueza e para outros continuidade ou atualidade de uma linha de pesquisa. Como condicionante, o fato do acesso não poder ser alcançado em detrimento da qualidade do conteúdo e da garantia da maior publicidade. Flexibilidade interpretativa é o que a literatura da área melhor retrata no momento, com várias perspectivas conceituais e modelos de desenvolvimento; e o fechamento e a estabilização dos RIs (p.e. um "modelo de negócio" ou um *design* dominante) é algo que aqueles que lutam bravamente para ganhar acesso à informação esperam que se realize em futuro próximo, e a seu favor.

Essa rápida apresentação é oportuna para que se tenha em mente que nas estratégias para desenvolvimento e implantação dos RIs é fundamental identificar os grupos sociais relevantes e os recursos e interesses que eles trazem para o processo; como eles se relacionam entre si, e com que agendas, implícitas e explícitas. Os alinhamentos de interesses podem unir um maior ou menor número de grupos sociais, em contraposição a outros, e antecipadamente orientar fortemente o curso da ação. Aqui entram também senão atores específicos (em geral, pesquisadores individuais com participação ativa) e políticas infraestruturais já estabelecidas, como aquelas relacionadas a *copyright* e propriedade intelectual.

Mais importante, e especialmente para os países em desenvolvimento, é fundamental reconhecer que as estratégias e recursos desses grupos sociais são contexto-dependentes, e que quaisquer experiências e *best practices* de outros paises, especialmente os desenvolvidos, informam, mas não podem ser tomadas como "tamanho único". Por exemplo, ainda que a comunidade de físicos de alta energia possam guardar, e de fato o faz, uma série de similaridades em todos os países do mundo no que diz respeito a sua prática de produção, acesso e uso de informação, eletrônica ou não, elas expressam também particula-

ridades locais, especialmente aquelas do Sul, e mais particularmente quando tentam se alinhar com as do Norte. Ou seja, alguns campos disciplinares são mais ou menos “internacionais”, mas todos refletem características locais. O mesmo ocorre para o caso das editoras científicas, comerciais ou não (e mesmo se elas existirem no Sul). No que diz aos interesses e estratégias das universidades/instituições de pesquisa, o quadro ainda é mais complexo, não só no que diz respeito à cultura local mas também no que diz respeito à forma como sua produção intelectual, e seus grupos de pesquisa, são financiados e avaliados.

Esse último ponto talvez seja um dos mais importante para as análises sobre RIs empreendidas na perspectiva do Sul, e mais especificamente no caso brasileiro: o sistema de avaliação acadêmica praticado no país. Se, como no caso do Brasil, as macropolíticas que orientam o sistema de recompensa dos pesquisadores e premia as instituições de pesquisa insistem em um modelo de publicação acadêmica exógeno (ou internacional), é pouco provável que o papel, a motivação e o alcance do potencial dos RIs possam se realizar da mesma forma como no Norte.

No que se segue, a lente do Programa SCOT é usada para empreender um rápida leitura de uma parte da literatura já disponível sobre RIs.

NATUREZA E FUNÇÃO DOS RIS

São duas as principais *rationales* que justificam o desenvolvimento e implantação dos RIs: por um lado, para tencionar (se não competir com) o modelo tradicional de publicação na ciência, e reduzir o poder de monopólio dos periódicos; por outro lado, para complementar (se não re-fundar) o sistema de comunicação acadêmica. No primeiro caso, o conceito de RI está intrinsecamente vinculado ao Movimento do Acesso Livre, e seu conteúdo, 100% de artigos que passaram pelo crivo dos pares, *preprints* e/ou postprints. No segundo caso, o RI apresenta-se mais como estratégia de gestão de conteúdos intelectuais de uma instituição de pesquisa; consequentemente, seu conteúdo é bastante diversificado, e sem a necessária vinculação com os pares. O que os

une, teoricamente, é uma infraestrutura tecnológica com padrão OAI, que permite o autoarquivamento e que possibilita a preservação digital. O acesso livre é uma premissa, não um atributo: alguns RIs operam com perfis de acesso diferenciados, e em alguns casos restringem o acesso ao texto completo, provavelmente por questões de *copyright*.

Como consequência, e também pela sua recenticidade, o conceito ainda é reconhecido como escorregadio. Bailey Jr. (2008), em um recente artigo de revisão, aponta para a polissemia do termo. Para alguns, um RI é um conjunto de serviços que uma universidade/instituição oferece para os membros da sua comunidade com vistas ao gerenciamento e disseminação de seu conteúdo intelectual digital (LYNCH, 2003); para outros, uma base de dados web de material acadêmico institucionalmente definido (WARE, 2004), ou ainda, como sendo constituído de uma coleção de itens capturados para preservar o capital intelectual de uma ou várias instituições (CROWN, 2002).

Guédon (2009) faz uma leitura muito instrutiva dessa polissemia. Aqueles que defendem um RI como um "serviço de informação", e colocam a ênfase no gerenciamento, organização e acesso, ecoam a partir de um contexto, de uma cultura e de uma prática de profissionais de informação, e buscam um caminho e um reposicionamento para as bibliotecas e seus profissionais em um ambiente crescentemente digital. A perspectiva de uma suposta neutralidade ao definir um RI como uma base de dados na *web*, com ênfase na interoperabilidade, vem dos esforços daqueles que operam nos limites entre as editoras científicas comerciais e as universidades, um "modelo de negócio" com séculos de sucesso. Por fim, aqueles que advogam que os RIs são uma estratégia para a reforma do sistema de publicação acadêmica estão profundamente envolvidos no Movimento do Livre Acesso, e mais particularmente, nas estratégias políticas para o acesso à produção científica oriunda das pesquisas em saúde financiadas pelo Estado, especialmente no contexto norte-americano.

Há ainda uma corrente que, sem necessariamente aderir a uma ou outra definição de RI, adotam uma perspectiva essencialmente funcional, e situam o RI como uma estratégia valiosa para dar visibilidade à produção acadêmica institucional, particularmente a partir

de um único ponto de acesso, o que contribuiria para aumentar o uso e impacto das pesquisas ali desenvolvidas (SWAN; CARR, 2008). Essa visão, também mais alinhada com o Movimento do Livre Acesso, mas não restritiva aos artigos revisados pelos pares, equaciona aumento de visibilidade com livre acesso, e aponta para o caráter mandatório de auto-arquivamento que emana das instituições. Aqui, o RI estaria também a serviço de uma estratégia de gestão institucional da pesquisa, como um guia para avaliação e, principalmente, para estimular e facilitar o re-uso dos conteúdos, acelerando o ciclo da pesquisa.

Uma visão representativa da própria comunidade científica é aquela defendida por Harnard (1990, 2006), um defensor incansável da extensão do sucesso do arXiv, de repositório temático de *preprints,* para repositório institucional. O foco aqui é no impacto e no uso da pesquisa, argumentando que muito se perde desse potencial dado o tempo e os custos da produção e distribuição. Ou seja, o depósito da versão final de um artigo científico em um RI (artigo esse já revisado por pares, e atendendo as questões de copyright e as políticas dos editores comerciais) é a forma mais rápida e efetiva de assegurar que um maior número de pesquisadores interessados possam lê-lo e citá-lo. Isso beneficiaria tanto o pesquisador individual como sua instituição, além de acelerar o ciclo da pesquisa. Essa visão é estritamente aderente ao Movimento do Livre Acesso e ao caráter mandatário do auto-arquivamento, e está em franca oposição àquelas estratégias de desenvolvimento de RIs que acolhem a diversidade de tipologia de conteúdos intelectuais produzidos pela instituição, o que, argumentam, seria muito mais a função de uma biblioteca digital.

Independente das diferentes visões quanto à função a ser desempenhada pelo RI, mesmos os entusiastas do primeiro minuto reconhecem que, até o presente, esses não vêm sendo usados em todo o seu potencial, e que as instituições têm falhado em suas estratégias de desenvolvimento, implantação e sustentabilidade. Para aqueles que veem o RI como espaço para artigos científicos revisados por pares, o grande desafio é, sem dúvida, a questão do *copyright* e a tradição de um "modelo de negócio" de publicação científica que vem se mostrando exitoso ao longo dos últimos séculos. Nessa configuração, as editoras

científicas representam o grupo social relevante com mais recursos e têm, com cerca tranquilidade, mantido seu papel de destaque, especialmente nos países em desenvolvimento. Caberia perguntar, especificamente no caso brasileiro, que outras estratégias e alternativas outros grupos sociais relevantes, como universidades/ institutos de pesquisa e órgãos governamentais, poderiam trazer para tensionar o modelo vigente.

Quando pensado como um serviço de informação, o êxito de um RI deve ser lido na perspectiva interna das instituições, ou, no âmbito das micropolíticas e das estratégias de gestão dos conteúdos digitais (LYNCH, 2003). No âmbito das micropolíticas, a palavra-chave é sustentabilidade, por exemplo: de investimentos financeiros e do comprometimento institucional, e estímulo à adesão e permanência dos principais atores, de gestão de conteúdo adequado. Aqui, de forma mais clara, estão envolvidas questões relacionadas ao envolvimento dos profissionais de informação e das bibliotecas, um grupo social extremamente relevante no que diz respeito à estruturação e gestão de conteúdo. A garantia do acesso é condição *sine qua non*, e questões de usabilidade e facilidade de interação são variáveis importantes. Para todos os casos, uma estruturação de conteúdos pobre e confusa e qualquer comprometimento no acesso aos conteúdos enfraquecem uma já tênue relação de confiança com o usuário, inerente ao uso de um "novo" serviço que, ainda na infância, precisa mostrar ao que veio.

Desenhado esse breve enquadramento de relações entre os grupos de interesse, o que se pode aprender com algumas experiências prévias?

ALGUMAS LIÇÕES DO CAMINHO JÁ TRILHADO

Williams e Lawton (2005) apontam que, no que diz respeito ao desenvolvimento e implantação de RIs, há extensa literatura sobre seus aspectos técnicos e operacionais, tanto no que diz respeito ao *hardware* e *softwares* disponíveis, como nas questões sobre arquitetura do sistema, interoperabilidade, metadados e preservação digital. Nessa perspecti-

va, as experiências relatadas situam um RI essencialmente como um serviço de informação e, exceto pela forma diferenciada de captura de material, guarda grande semelhança com a literatura voltada ao desenvolvimento de sistemas de recuperação de informação.

No geral, a literatura retrata que são várias as demandas para desenvolvimento de RI, e que as mesmas continuam em evolução, e que não há nenhum roteiro/protocolo padrão indicando a melhor maneira de progredir. Sheereves e Cragin (2008) reconhecem que, vencida quase uma década após os primeiros RIs serem implantados no Reino Unido e nos EUA, a motivação que os guiou ainda não é clara. Antes, para, e, em cada país, uma demanda diferente, uma configuração de forças, recursos e interesses entre os grupos sociais relevantes específica, e um curso de ação particular.

Na perspectiva micro das instituições de pesquisa, ou, no espaço onde o RI deixa de ser conceito para ser operativo, de comum acordo o fato de ser um trabalho extremamente técnico, exigindo a concepção e implementação de infraestrutura tecnológica e de informação sólidas, mas também uma tarefa que exige um aporte gerencial de excelência, com planejamento contínuo, priorizando e coordenando as ações com as expectativas dos diversos grupos sociais relevantes.

No que diz respeito ao grau de sucesso dos repositórios institucionais, as evidências apontam ainda que é algo irregular. Em 2005, Lynch e Lippincott, (2005) promoveram um inquérito com universidades e faculdades de artes liberais nos Estados Unidos. Ali, 40% das universidades já contavam com RIs operacionais. DSpace foi o pacote de gerenciamento de conteúdo predominante listado pelos entrevistados. O tamanho descrito destes RIs variou entre centenas de milhares de objetos digitais (acima de 10 terabytes de espaço) para menos de uma dúzia (inferior a um gigabyte), embora houvesse confusão entre os entrevistados sobre o que seria um "objeto digital" – um banco de dados, ou um registro de informação.

A tipologia dos materiais armazenados era diversificada, incluindo *e-prints*, teses e dissertações eletrônicas, coleções especiais digitalizadas, materiais multimídia, materiais de cursos, dentre outros. De forma clara, essa variedade de tipologias está diretamente ligada à motiva-

ção que guiou a implantação do RI, o que, por outro lado, cria seus próprios problemas. Em sua pesquisa, Macdowell (2007) registra que somente cerca de 13% dos conteúdos dos RIs passaram por revisão pelos pares. Argumenta-se que essa diversidade, ou inconsistência de coleções, pouco contribui para quesitos de usabilidade e confiança do usuário, além de serem iniciativas que se distanciam do apelo central do Movimento de Livre Acesso (POYNDER, 2006).

A participação do corpo docente da instituição foi, para todos os entrevistados, apontada como voluntária, ou seja, sem caráter mandatário. No geral, é uma participação considerada muito baixa.

Sem surpresas, Lynch e Lippincott, (2005) apontam que alcançar uma taxa significativa de participação (o que quer que seja isso) é o principal indicador de sucesso dos RIs citado na literatura. Alguns gestores de RIs tentam recrutar a contribuição dos profissionais enfatizando a importância da preservação digital, e mesmo como aferição do uso dos resultados das pesquisas (DAY, 2004). Mas os fatores que motivam os acadêmicos são, de fato, mais complexos. Análises recentes (SWAN et al., 2005; FOSTER; GIBBONS, 2005; KENNAN; WILSON, 2006) confirmam as suspeitas de que o comprometimento com a visibilidade institucional e/ou a possibilidade de uma avaliação da produtividade científica na perspectiva institucional interna muito certamente não são os fatores motivadores para participar do empreendimento RI. Ao contrário, a motivação parece emergir fortemente do desejo de reconhecimento pessoal e do impacto entre os pares.

Uma pesquisa similar, também em 2005, foi realizada em universidades de dez países europeus - Bélgica, França, Reino Unido, Dinamarca, Noruega, Suécia, Finlândia, Alemanha, Itália e Holanda -, bem como no Canadá e Austrália (VAN WESTRIENEN; LYNCH, 2005). O número de repositórios institucionais em cada país variou de no mínimo 1,5% das universidades (caso da Finlândia, com uma única universidade), para no máximo 100%, na Alemanha, Noruega e Holanda. Aqui também a tipologia de conteúdos restringiu-se quase exclusivamente a publicações do corpo docente.

Como o estudo americano, o inquérito europeu também identificou a baixa participação do corpo docente no armazenamento de objetos

em seus repositórios. Van Westrienen e Lynch (2005) identificaram várias razões para essa não participação:

- Dificuldades para informá-los e convencê-los a participar;
- Confusão e incertezas a respeito de questões sobre propriedade intelectual;
- Crédito acadêmico e uso do material estocado (receio de plágio, por exemplo);
- A percepção de que os conteúdos de livre acesso sejam de baixa qualidade, e
- A falta de políticas obrigatórias para depósito.

Bailey Jr. (2006) também conduziu um inquérito em 2005, na perspectiva dos gestores de 123 bibliotecas da Association of Research Library (ARL), EUA. Aqui, as bibliotecas assumiram a responsabilidade de desenvolver e implantar RIs em resposta à demanda institucional. No geral, esses profissionais de informação foram, de longe, os atores mais ativos nesse empreendimento, com uma participação também ativa dos profissionais de TI. As principais motivações para o desenvolvimento de RI citadas foram, em ordem de importância: a promoção da visibilidade institucional, a preservação, o livre acesso e a administração dos conteúdos intelectuais da instituição. Um RI típico possui, em média, uma coleção de cerca de 3.800 objetos digitais, de tipologia diversificada, de *preprints* a *postprints*, relatórios técnicos, material multimídia, dentre outros. Grande parte dos RIs possuem políticas (normas/procedimentos) explícitas que orientam, entre outros pontos, o perfil de acesso dos diferentes usuários, o que significa que nem todos praticam o acesso livre em sua plenitude. O auto-arquivamento também era submetido a um perfil de autorização, e no geral, já eram depositados tendo passado pelo crivo das comunidades que definiam o perfil nas coleções. Enquanto a totalidade dos RIs almejam a preservação dos objetos digitais, poucos ainda a praticavam. Raras eram ainda as avaliações de uso e impactos dos RIs, o que explicita o pouco cuidado com a visão dos usuários.Novamente, o ponto nevrálgico apontado na pesquisa diz respeito às estratégias para recrutamento de material a ser depositado.

Palmer e Newton (2008), por fim, apresentam uma análise das abordagens de desenvolvimento adotadas por três bibliotecas universitárias norte-americanas, onde procurou comparar as escolhas, as estratégias e condições para atividades direcionadas ao desenvolvimento.

O estudo aponta que as diferenças mais acentuadas resultam de iniciativas como o equilíbrio, a aquisição de conteúdo e a prestação de serviços. A preocupação com propriedade intelectual é item predominante, e os objetivos e políticas do repositório são muitas vezes implícitas, e o valor do depósito para o corpo docente e para a universidade emergem de várias formas. O planejamento complexo, a gestão e o trabalho técnico dos desenvolvedores de repositório dependem cada vez mais de uma coordenação com os bibliotecários e a interação existente entre eles e os acadêmicos/pesquisadores.

Três abordagens foram identificadas em cada uma das três instituições analisadas. A instituição A distingue-se por uma forte ênfase em serviços, e tem uma abordagem pro-ativa ao trabalhar com pesquisadores para resolver seus problemas de gestão da informação e de curadoria de dados. A instituição B apresenta um equilíbrio de desenvolvimento de conteúdo orientado politicamente e inclui iniciativas de serviços seletivos que têm implicações para além do campus. Na instituição C, os desenvolvedores concentraram-se na construção de uma base forte de conteúdo, confrontando-se diretamente com desafios de propriedade intelectual.

Os três casos sugerem um leque de respostas produtivas para os muitos desafios que enfrentam os desenvolvedores de repositórios institucionais, à medida que os repositórios amadurecem, expandem-se e integram-se mais com as atividades das bibliotecas, e continuam a sua importante contribuição para o sempre mutante empreendimento da comunicação científica.

Dentro do tema mais amplo de equilíbrio de conteúdo e demandas de serviços, três dimensões de desenvolvimento de RIs mostraram-se importantes nas entrevistas: objetivos e política, propriedade intelectual e reconhecimento de valor.

Medir o sucesso de um repositório não é somente uma questão de números. Cabe antes perguntar quais eram as metas e se as mesmas

foram alcançadas. O perfil da instituição e as áreas disciplinares/comunidades de pesquisa que ela acolhe serão os guias para o escopo do RI, e para o quantitativo (e qualitativo) potencial que se pode alcançar. Se o comprometimento dos acadêmicos é fundamental, vários são os fatores que os levam a participar, e esses precisam ser melhor identificados e analisados. A sinergia com as atividades e estratégias já existentes, especialmente com os mecanismos de avaliação existentes, é fator fundamental para uma maior participação.

Ferreira e colaboradores (2008), no relato da política de desenvolvimento do RI da Universidade do Minho, Portugal, talvez reconhecidamente um *case* de sucesso internacional, comparam a experiência aos trabalhos de Hércules, ainda que não somente pelos aspectos técnicos ou garantia de financiamento institucional, mas principalmente pela forma com que os RIs interferem nas práticas tradicionais dos acadêmicos. E deixam aos leitores uma palavra de esperança: "[...] Nevertheless, as soon as an institutional repository is set up, all of the academy's research output is expected to be placed in the repository in order to increase the academy's visibility, usage and impact (among other things, such as constituting the long-term memory of the academy)." Talvez a grande diferencial da Universidade do Minho tenha sido a inteligência gerencial para desenvolver valor agregado no RI e prover formas criativas e legais para ter acesso a material com restrição de uso.

O que essas experiências deixam claro é que, a despeito do maior ou menor grau de sucesso dos RIs, os profissionais de informação, e as bibliotecas, são o grupo social relevante mais importante nessa configuração, e podem assim ocupar um papel de destaque. A proatividade desses profissionais pode, por exemplo, forjar alianças criativas com os gestores e acadêmicos/pesquisadores, levando a novas questões e novas soluções. Engajar os pesquisadores na primeira fase de concepção e desenvolvimento dos RIs é fundamental. Guédon (2009) observa que as demandas dos mesmos, enquanto leitores/usuários dos RIs, parecem estar sendo negligenciadas, ou, mal interpretadas.

Cabe, no caso brasileiro, perguntar se os profissionais de informação (obviamente aqui incluídos os bibliotecários) que estão ligados ao serviço nas universidades e instituições de ensino e pesquisa, estão

mobilizados e comprometidos com esse novo desafio. Quiçá a resposta seja positiva!

ORIENTAÇÕES PARA O FUTURO: UMA RESPOSTA PARA CADA PERGUNTA

Um leitor atento vai perceber que pouco se apresentou, ao longo do texto, uma perspectiva que venha das "macropolíticas" públicas, onde experiências como a norte-americana (capitaneada pela National Institutes of Health – NIH), a britânica (que nasceu no privado com a *Wellcome Trust* e se institucionalizou no Estado) e a australiana (com o programa governamental Asher) são experiências que, por certo, devem servir de aprendizado para outros países. Aqui, duas justificativas são dignas de nota.

Primeira, a ausência desse olhar das macropolíticas decorre da própria perspectiva metodológica adotada para fazer uma leitura dos RIs – no programa SCOT, como em todo conjunto de abordagens "construtivistas" da ciência e tecnologia, não há um *out there,* uma configuração política macro tomada como dada. Cada grupo social relevante, com seus recursos e seus interesses, traz embutido em si uma leitura das mesmas. E é no seio da concepção e do desenvolvimento das novas tecnologias, na resistência e na diferença, que as políticas, macro e micro, são forjadas. Políticas são consequências, e não causas.

Segunda, há ainda um grande distanciamento entre as dimensões micro (o "colocar a mão na massa" e sua política cotidiana) e macro (orientações que emanam de níveis superiores, como de órgãos de financiamento) nas discussões sobre os RIs, novamente reflexo de suas diferentes concepções vigentes. Enquanto serviço de informação, as discussões estão muito mais orientadas para o âmbito interno das instituições; enquanto uma estratégia do Movimento do Livre Acesso, as discussões estão mais orientadas para o âmbito das políticas públicas. Essa flexibilidade interpretativa ainda não permite falar sobre os RIs em um *continuum* que vai da macro para micropolítica, e vice-versa. Fazer a defesa dos repositórios na perspectiva do Movimento

de Livre Acesso é uma estratégia *up-down*; desenvolver repositórios como serviços é uma estratégia *down-up*. Em algum momento, elas haverão de se encontrar.

Das citadas experiências de sucesso citadas acima, de três países diferentes, e descritas com cuidado por Guédon (2009), são pinçadas algumas conclusões, na perspectiva das macropolíticas: o envolvimento e a decisão firme do Estado, especialmente por meio de seu corpo legislativo, é fundamental. As agências de fomento parecem ser o fiel da balança, e são elas que melhor encarnam a legitimidade do caráter mandatório de depósito de artigos revisados por pares, tornando-se o melhor contraponto ao poder de fogo das editoras científicas comerciais.

O que essas experiências deixam antever ainda é que um setor de financiamento de pesquisa com perfil mais público que privado é uma vantagem, obviamente porque aí o Estado ganha mais independência em suas decisões. Mas o que é uma vantagem pode também ser uma desvantagem: se o Estado não está mobilizado ou tem pouca capacidade de escuta, os desafios são maiores.

Emerge também dessas experiências, de forma clara, o papel fundamental desempenhado pelos profissionais de informação, na figura de suas sociedades/organizações de classe. Talvez seja isso um testemunho do reconhecimento e do respeito que os mesmos despertam em gestores e pesquisadores, fruto de décadas de trabalho de excelência, especialmente no campo da pesquisa em saúde.

De todas as lições, a mais importante, e também a mais óbvia: a plena realização do potencial dos RIs, enquanto ampliação do acesso a artigos científicos revisados por pares, depende fundamentalmente de sua integração/harmonização com os mecanismos de avaliação científica em curso. Enquanto essas ligações não existirem, ou forem frágeis, qualquer estratégia de desenvolvimento será infrutífera, se não irrelevante.

Somente assim os RIs poderão ser a resposta para várias perguntas!

REFERÊNCIAS

BAILEY JR, C. **Institutional repositories, tout de suite**. 2008. 10 p. Disponível em : http://digital-scholarship.org/ts/irtoutsuite.pdf. Acesso em 25 nov. 2009

BOHLIN, I. Communication regimes in competition:the current transition in scholarlycommunication seen through the lens of the sociology of technology. **Social Studies of Science,** v. 34, n. 3, p. 365-391, June 2004.

BORGMAN, C. L. **Scholarship in the digital age**: nformation, infrastructure, and the Internet. London: The MIT Press, 2007.

CROW, R. **The case for institutional repositories**: a SPARC position paper. SPARC: Washington. 2002. Disponível em: <http://www.arl.org/sparc/IR/ir.html>. Acesso em: 15 set. 2003.

DAY, M. **Institutional repositories and research assessment**: a supporting study for the ePrints UK Project. 2004. Disponível em: <http://eprints-uk.rdn.ac.uk/project/docs/studies/rae/rae-study.pdf>. Acesso em 30 nov. 2009

FERREIRA, M. et al. Carrots and sticks: some ideas on how to create a successful institutional repository. **D-Lib Magazine**, v. 14, n. 1/2, Jan./Feb.2008. Disponível em: <http://webdoc.sub.gwdg.de/edoc/aw/d-lib/dlib/january08/ferreira/01ferreira.html> Acesso em: 15 dez. 2009.

FOSTER, N. F. ; GIBBONS, S. Understanding faculty to improve content recruitment for institutional repositories. **D-Lib Magazine**, v. 11, n. 1, Jan. 2005. Disponível em: <http://dlib.org/dlib/january05/foster/01foster.html>. Acesso em: 19 dez. 2009.

GUEDÓN, J. C. It´s a repository, it´s a depository, it´s an archive...: open access, digital collections and value. **Arbor, ciencia, pensamiento y cultura**, v. 185, n. 737, p. 581-595, mayo-jun. 2009. Disponível em: < http://arbor.revistas.csic.es/index.php/arbor/ article/viewFile/315/316>. Acesso em: 20 dez. 2009

GUIMARÃES, M. C. S. Entre o superlativo e o diminutivo: as tecnologias a serviço da difusão da informação científica. In: MANDARINO, A. C. de S.; GOMBERG, E. **Leituras de novas tecnologias e saúde**. Salvador: EDUFBA, 2009. p. 117-137.

HALLORAN, J. D. Information is the answer, but what is the question? **Journal of Information Science**, v.7, n. 4-5, p.159-167, 1983.

HARNAD, S. Scholarly skywriting and the prepublication continuum of scientific inquiry. **Psychological Science**, *v. 1, p. 342-343, 1990. reprinted in Current Contents v. 45, p. 9-13, November 11 1991. Disponível em:* <http://cogprints.org/1581/0/harnad90.skywriting.html>*. Acesso em: 16 dez. 2009.*

___. Publish or perish – self-archive to flourish: the green route to open access. **ERCIM News 64**, 2006. Disponível em: <http://users.ecs.soton.ac.uk/harnad/Temp/harnad-ercim.pdf>. Acesso em: 15 dez. 2009.

KENNAN, M. A.; WILSON, C. Institutional repositories: review and an information systems perspective . **Library management**, v. 27, n. 4/5, p. 236-248, 2006. Disponível em: <http://unsworks.unsw.edu.au/vital/access/manager/Repository/unsworks:24>. Acesso em: 15 dez. 2009.

KING, C. et al. **Scholarly communication**: academic values and sustainable models. UC Berkeley: Center for Studies in Higher Education, 2006. Disponível em: <http://www.escholarship.org/uc/item/4j89c3f7>. Acesso em: 16 dez. 2009

KLING, R . Social informatics: a new perspective on social research about information and communication technologies. **Prometheus**, v. 18, n. 3, p. 245-266, 2000.

___. Learning about information technologies and social change: the contribution of social informatics. **The Information Society**, n. 16, p. 217-232, 2000.

___. What is social informatics and why does it matter? **The Information Society**, v. 23, n. 4, p. 205-220, July, 2007.

KURAMOTO, H. Informação científica: proposta de um novo modelo para o Brasil. **Ci. Inf.**, Brasília, v. 35, n. 2, Aug. 2006 . Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652006000200010&lng=en&nrm=iso>. Acesso em: 5 dez. 2009. doi: 10.1590/S0100-19652006000200010.

LYNCH, C. A. Institutional repositories: essential infrastructure for scholarship in the digital age. **ARL Bimonthly Report,** 2003. Disponível em: <http://www.arl.org/newsltr/226/ir.html>. Acesso em: 15 dez. 2009.

LYNCH, C. A.; LIPPINCOTT, J. K. Institutional repository deployment in the United States as of early 2005. **D-Lib Magazine**, v.11, n. 9, Sept. 2005. Disponível em: <http://www.dlib.org/dlib/september05/lynch/09lynch.html>. Acesso em: 20 dez. 2009.

MACDOWELL, C. S. Evaluating institutional repository development in American academe since early 2005: repository by the numbers, part 2. **D-lib Magazine**, v. 13, n. 9/10 , Sept./Oct. 2007. Disponível em: < http://www.dlib.org/dlib/september07/mcdowell/09mcdowell.html>. Acesso em: 18 dez. 2009.

OPEN SOCIETY INSTITUTE. **Open access publishing and scholarly societies**: a guide. New York: Open Society Institute, 2005. Disponível em: < www.soros.org/openaccess/scholarly_guide.shtml> Acesso em: 16 dez. 2009.

PALMER, L. C. T.; NEWTON, M. P. Strategies for institutional repository development: a case study of three evolving initiatives. **Library Trends**, v. 57, n. 2, p. 142-67, 2008 (special issue). Disponível em: http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1131&context=lib_research. Acesso em: 10 jan. 2010.

PINCH, T.; BIJKER, W. The social construction of facts and artifacts or how the sociology of science and the sociology of technology might benefit each

other. In: BIJKER, W.; HUGHES,T.; PINCH,T. 1984. **The social construction of technological systems**: new directions in the sociology and history of technology. Cambridge, MA: MIT Press, 1987.

POYNDER, R. **Clear blue water**. (Eprint insiders) 7 Mar. 2006 Disponível em: http://www.eprints.org/community/blog/index.php?/archives/43-Clear-blue-water.html >. Acesso em: 20 dez. 2009.

SHEEREVES, S ; CRAGIN, M. H. Institutional repositories: current state and future. **Library Trends**, v.57, n.2, p. 89-97, Fall, 2008.

STIGLITZ, J. Knowledge as a global public good. In: KAUL, I.; GRUNBERG, I.; STERN, M. A. (Ed.). **Global public goods**: international co-operation in the 21st century. New York: Oxford University, 1999. p. 308-326.

SWAN, A. et al. Developing a model for e-prints and open access journal content in UK further and higher education. **Learned Publishing**, *v.* 18, n. 1, p. 25-40, 2005.

SWAN, A.; CARR, L. Institutions, their repositories and the Web. **Serials Review**, v. 34, n. 1, 2008. Disponível em: http://eprints.ecs.soton.ac.uk/14965/. Acesso em: 27 nov. 2009.

VAN WESTRIENEN, G.; LYNCH, C. A. Academic institutional repositories: deployment status in 13 nations as of mid 2005. **D-Lib Magazine**, v. 11, 2005. Disponível em: <doi:10.1045/september2005-westrienen>. Acesso em: 17 dez. 2009.

WARE, M. **Pathfinder research on web-based repositories**. London: Publisher and Library/Learning Solutions, 2004. Disponível em: < http://www.palsgroup.org.uk/palsweb/

palsweb.nsf/79b0d164e01a6cb880256ae0004a0e34/8c43ce800a9c67cd80256e370051e88a/$FILE/PALS%20report%20on%20Institutional%20Repositories.pdf. Acesso em: 15 dez. 2009.

WILLIAMS, R. The Social shaping of information and communications technologies. In: KUBICECK, H.; DUTTON, W. H.; WILLIAMS, R. **The social shaping of information superhighways**. Frankfurt: Campus Verlag; New York: St Martin Press, 1997.

WILLIAMS, R.; EDGE, D. The social shaping of technology. **Research Policy**, 25, p. 865-899, 1996.

WILSON, M.; HOWCROFT, D. Re-conceptualising failure: social shaping meets IS research **European Journal of Information Systems**, v. 11, n. 4, p. 236-250, Dec. 2002.

Repositórios institucionais confiáveis:

Repositório institucional como ferramenta para a preservação digital

Cristiane Yanase Hirabara de Castro
Marcos Sfair Sunye
Luiz Carlos Erpen de Bona
Marcos Alexandre Castilho

INTRODUÇÃO

A informação tornou-se a ferramenta mais utilizada pelas empresas e pelas organizações institucionais publicas ou privadas para a descoberta e divulgação de conhecimento. A característica dinâmica dessas informações facilitadas por recursos tecnológicos disponíveis impede que essas informações permaneçam fisicamente em unidades literárias ou documentais, como as bibliotecas tradicionais que se dedicavam a controlar informações e documentos físicos e não digitais. Desta forma, as organizações e a população em geral estão transformando e produzindo documentos exclusivamente em formato digital: textos, grandes bases de dados e imagens tornando evidente que o contexto mundial já apresenta uma característica fortemente dependente de documentos digitais como um meio para registrar e preservar as produções individuais, institucionais e governamentais.

A informação em formato digital, porém, está sujeita a inúmeras intervenções internas e externas como: perda, adulteração e destruição, degradação física, obsolescência tecnológica de *hardware*, *software* e formatos, entre outros fatores que podem modificar o seu conteúdo, comprometendo sua qualidade e integridade (RAMOS et al, 2004). É necessário, portanto, que os arquivos facilitem o estabelecimento

de políticas, procedimentos e práticas para assistir às organizações e apoiá-las a criarem e manterem documentos fidedignos, autênticos, acessíveis e preserváveis.

Os Sistemas de Preservação Digital (SPD) surgiram para atender a demanda na acessibilidade e segurança de documentos digitais no longo prazo. Em geral um SPD envolve a implementação de três processos: a aquisição, a indexação e a distribuição do material a ser preservado. A aquisição de um material envolve basicamente duas fases: a escolha da forma de armazenamento e a determinação do formato e local de armazenamento visando a preservação desse material no longo prazo. O processo de indexação automática deve permitir a classificação do material obtido e prover o armazenamento de informações relevantes ao processo de recuperação e distribuição.

No entanto, a preservação digital não consiste apenas na forma de armazenamento, mas na capacidade de garantir que a informação digital permaneça acessível e com qualidades de autenticidade suficientes para que possa ser interpretada no futuro recorrendo-se a uma plataforma tecnológica diferente da utilizada no momento da sua criação. Sobretudo, a preservação digital é a atividade responsável por garantir que a comunicação seja possível não só através do espaço, mas também através do tempo (FERREIRA, 2006).

Assim, cabe ao Arquivamento Digital (AD) a responsabilidade pela integridade dos dados armazenados. Dentre os objetivos do AD estão: a garantia de obter melhor desempenho no acesso aos documentos em relação ao arquivamento tradicional e garantir também o arquivamento a longo prazo utilizando dispositivos com vida útil, atualmente, em torno de 5 anos.

Ao ressaltar a importância do processo de preservação para as bibliotecas digitais é necessário ter uma ideia mais concreta dos componentes e processos relacionados à preservação digital e dos elementos que o compõe. A seção 2 introduz estes conceitos. Em seguida apresentamos os resultados obtidos nas pesquisas realizadas pelo centro de Computação Científica e *software* Livre da Universidade Federal do Paraná (C3SL/UFPR), em que o foco são os modelos de arquitetura para arquivamento digital em ambientes distribuídos.

Finalmente descrevemos os serviços relacionados a um AD e como eles se inserem em uma rede de Bibliotecas Digitais.

ELEMENTOS DO ARQUIVAMENTO DIGITAL

Em 2003, a ISO estabeleceu algumas normas para o arquivamento digital (ISO 1472:2003), através da iniciativa Open Archival Information System (OAIS), para identificar e padronizar os principais componentes funcionais e objetos de informação num sistema de arquivos com pretensões no longo prazo. E desde então o OAIS é considerado o modelo de referência para a preservação digital.

O objetivo do modelo OAIS é suportar e garantir que as características mais relevantes do arquivamento digital como a disponibilidade, durabilidade e confiabilidade dos dados sejam mantidos e considerados pelos SPD, por meio de padrões que permitem a manutenção, o compartilhamento e a distribuição do material preservado.

A disponibilidade do material está relacionada à acessibilidade do material digital, ao controle e permissão de acesso e ao formato compatível por meio da evolução de versões e distribuição. Confiabilidade consiste em conservar o material intacto no longo do tempo e auditorias frequentes para verificar a existência de obsolescência do material ou arquivos corrompidos.

Um dos maiores problemas na implementação de SPD é garantir que os dados não se tornem obsoletos devido a evolução dos meios de armazenamento físico ou dos *softwares* em que foram produzidos. Outros problemas também devem ser considerados como: falhas de mídia, hardware ou *software*, erros de comunicação e falhas na estrutura de rede, ataques internos ou externos, desastres naturais e erros humanos, dentre outros.

Ainda nesse contexto, considerando o modo acelerado com que a quantidade de informação na Internet cresce, existem também problemas relacionados diretamente a computação e armazenamento: o crescimento acelerado de informação requer servidores não convencionais, suporte de hardware específico para suportar o armazenamento

de grande quantidade de informação e permitir que a escalabilidade dos equipamentos acompanhe o crescimento de informação (geradores, nobreaks, arrays de disco, entre outros) e recurso humano especializado.

Tecnicamente, o arquivamento digital objetiva a preservação de objetos digitais. Um objeto digital, segundo Hunter e Choudhury (2004) e Lorie (2002) é definido como todo e qualquer objeto de informação que possa ser representado através de uma sequência de dígitos binários. São exemplos de objetos digitais os documentos de texto, fotografias, bases de dados, entre outros materiais produzidos digitalmente.

O processo de preservação de objetos digitais está relacionado às diferentes formas com que esses objetos podem ser armazenados.

Segundo Lee, Slatery e McCary (2002) as estratégias de preservação digital existentes são três: a emulação, a migração e encapsulamento. Mas para Thibodeau (2002) as estratégias de preservação não podem ser resumidas apenas em emulação, migração e encapsulamento, ele propõe um mapa bidimensional, posicionando no seu extremo esquerdo as estratégias centradas na preservação do objeto físico/lógico e no extremo oposto as estratégias centradas na preservação do objeto. Essas estratégias vão desde a preservação de tecnologia, que corresponde à conservação e manutenção de todo *hardware* e *software* necessários à correta apresentação do objeto lógico até as técnicas mais utilizadas, como a migração. A migração ou conversão é o processo de transferência periódica de material digital de uma dada configuração de *hardware/software* para uma outra, ou de uma geração de tecnologia para outra subsequente.

A escolha da melhor estratégia de preservação a ser utilizada depende do grau de obsolescência do objeto. E embora cada estratégia apresente características distintas, elas podem ser combinadas umas as outras para serem utilizadas, dependendo da necessidade e do contexto em que se encontram os objetos a serem preservados.

À medida que as estratégias de preservação foram se consolidando e as técnicas foram sendo aperfeiçoadas surgiu a necessidade de uma padronização para o processo de preservação digital.

Estudos foram iniciados na década de 1990, pelo Consultative Comitee for Space Data Systems (CCSDS) em conjunto com a International

Organization for Standardization (ISO), mostrando a preocupação de profissionais da área. Os esforços desses profissionais foram consolidados com o surgimento do modelo padronizado Open Archival Information System (OAIS) para o processo de preservação digital.

MODELO DE REFERÊNCIA OAIS

O modelo OAIS descreve as interfaces internas e externas do sistema e os objetos de informação que são manipulados no seu interior. A figura 1 ilustra o modelo OAIS, seus componentes e o processo de preservação, desde a submissão ao acesso das informações. Cinco componentes são descritos nesse modelo: planejamento de preservação, ingestão, gestão dos dados, acesso e administração. Produtor e Consumidor são as entidades externas responsáveis pela submissão e acesso da informação preservada Lavoie (2004).

O componente de Planejamento de Preservação é responsável por definir as estratégias de preservação mais adequadas ao material de informação submetido e fazer o monitoramento de todo processo de preservação.

O material submetido, pelo Produtor, é denominado Pacote de Informação de Submissão (PIS). O componente de Ingestão é a interface entre entidade externa e o sistema de preservação que permite o processo de incorporação do PIS pelo Repositório de Dados e pela produção da Informação Descritiva. A metainformação, produzida no processo de ingestão deve conter informações sobre a localização e descoberta do material de submissão

O componente Repositório de dados é responsável por manter a integridade do PIS e armazenar o material de informação, descrito na figura por Pacote de Informação de Arquivo (PIA) Lavoie (2004).

O componente Acesso é a interface entre a entidade externa e o sistema de preservação, que possibilita o acesso do material preservado pelo consumidor. Este componente é responsável pela recuperação da metainformação e do material preservado, agora denominado Pacote de Informação de Disseminação (PID) que será consumido pela entidade externa.

O PID e o PIA podem não ser idênticos, dependendo das estratégias de preservação utilizadas e do tipo de informação requisitada pelo consumidor, isto é, a necessidade do consumidor pode se restringir ao acesso de apenas um subconjunto de informação do material preservado e não de todo PIA.

O componente Administração é responsável por gerenciar e monitorar todo o processo de preservação e também dos artefatos gerados e consumidos por cada componente. Este componente é responsável também pela manutenção do Repositório de Dados e da Gestão de Dados.

Figura 1 - Modelo de referência OAIS
Fonte: CCSDS (2002).

Baseado no modelo de referência OAIS, destaca-se o projeto OAI (Open Archives Initiative) por fornecer vantagens como a cooperação e compatibilidade entre bibliotecas digitais. A arquitetura do protocolo OAI propõe a separação entre metadados e objetos digitais.

A Iniciativa Open Archives (OAI) desenvolve e promove padrões de interoperabilidade, com o objetivo de facilitar a disseminação eficiente de conteúdo (CCSDS, 2002). Uma federação OAI é claramente baseada na separação entre os Provedores de Dados (Data Providers – DP) e os Provedores de Servi¸co (Service Providers – SP), conforme pode ser observado na arquitetura apresentada na figura 2.

Os Provedores de Dados são aplicações que possuem a responsabilidade de gerenciar os repositórios de conteúdo, onde os objetos digitais (artigos, dissertações, teses, entre outros) são armazenados. Além disso, DPs implementam o protocolo OAI-PMH com o intuito de criar e expor os metadados (informações descritivas sobre os registros dos repositórios) de seus objetos (ROTHENBERG, 1999).

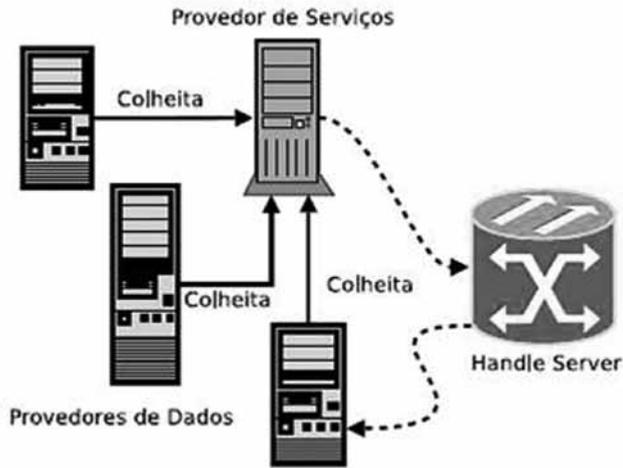



Figura 2 - Arquitetura de uma federação OAI

Provedores de Dados podem ser concebidos com suporte nativo ao protocolo OAIPMH – através da utilização de *softwares* desenvolvidos para apoiar a Iniciativa Open Archives, como DSpace[1] e Eprints[2] – ou ser criados sobre repositórios de dados já existentes, através da implementação de uma interface com o protocolo, como é o caso do arXiv[3] e SciELO[4].

A inserção de um objeto digital no DP é realizada em três passos: (i) um metadado é criado pelo DP seguindo o padrão Dublin Core; (ii) um

1 http://www.dspace.cam.ac.uk/

2 www.eprints.org/

3 http://arxiv.org/

4 http://scielo.org

Identificador Permanente de Objeto (Digital Object Identifier - DOI) é criado e inserido no metadado; e (iii) o *upload* do objeto é realizado e seu conteúdo armazenado no repositório. (BECKER et al., 2008)

Uma vez que os objetos digitais estão armazenados nos repositórios e seus respectivos metadados disponíveis, surge a figura dos Provedores de Serviço, que são responsáveis por realizar a "colheita" (*harvesting*) de todos os metadados em um conjunto de DPs, respeitando os critérios de coleta disponíveis no protocolo OAI-PMH. Após colhidos, os metadados são armazenados localmente, organizados e utilizados como base para disponibilizar um mecanismo de busca unificado sobre os registros dos repositórios. Desta forma, ao acessar um SP e realizar uma busca, o usuário é direcionado ao DP responsável pelo objeto.

Neste sentido, o objetivo dos trabalhos que estão sendo desenvolvidos na UFPR concentra-se em conceber arquiteturas para preservação digital baseada em redes P2P, que possibilite a criação de sistemas de baixo custo e que aumente a confiabilidade e a disponibilidade dos objetos em Bibliotecas Digitais OAI.

Na seção 3 serão relatadas as experiências da UFPR com a preservação digital e redes *peer-to-peer* (P2P) juntamente com alguns conceitos técnicos necessários para a melhor compreensão do leitor acerca dos modelos de arquiteturas para ambientes distribuídos resultantes das pesquisas realizadas no Departamento de Informática (DI).

ARQUIVAMENTO DIGITAL BASEADO EM REDES P2P

O custo para implementar um SPD que conserve as características desejáveis de um AD eficiente e ainda resolva os problemas, descritos na seção 2, é muito alto, isso se deve ao alto valor do suporte de *hardware*, a adaptação do ambiente físico (instalação de rede e elétrica) e os recursos humanos que além de caros não são encontrados facilmente no mercado. Assim, a criação de novas abordagens para preservação digital deve contar com sistemas de baixo custo que permitam replicar grandes volumes de dados, mecanismos de auditoria para verificar a

integridade dos mesmos além de mecanismos de recuperação que garantam a sua disponibilidade.

A replicação pode ser uma solução viável que aumenta a confiabilidade, a disponibilidade, a tolerância a falhas (de *hardware* e *software*) e o desempenho do sistema, sem considerar que o custo de hardware e *software* para isso é relativamente baixo. A grande desvantagem da replicação é que a criação de cópias aumenta o uso dos dispositivos de armazenamento e dificulta a segurança e a atualização dos dados.

Além disso, as ferramentas existentes assumem que todos os objetos do sistema possuem a mesma importância, do ponto de vista de preservação. Tendo em vista que, em sua grande maioria, soluções de preservação digital são baseadas em replicação, é importante identificar os objetos que devem ser replicados com maior prioridade, ou seja, possuir mais réplicas no sistema.

Considerar uma única estratégia de replicação para todos os objetos cria um grande número de réplicas desnecessárias no sistema, uma vez que não há bom aproveitamento do espaço em disco. Assim, de maneira geral, tal solução impossibilita a otimização da redução de custo com relação à mídias de armazenamento.

O uso do paradigma P2P para o desenvolvimento de SPD, apesar de ser uma implementação potencialmente mais complexa, gera vantagens na distribuição de carga, visto que todas as máquinas (*peers*) dividem as tarefas do servidor colaborando entre si e compartilhando recursos ociosos de computação e armazenamento. Outras vantagens são oferecidas por esse paradigma: P2P utiliza a filosofia de multiplicar as cópias para aumentar a disponibilidade, desempenho de download etc. A replicação, possui uma arquitetura já consolidada, chamada Distributed Hash Table (DHT), apresenta um grande aumento na escalabilidade uma vez que a capacidade do sistema aumenta proporcionalmente com o número de *peers* participantes, são auto-organizáveis ou seja, o sistema se organiza sem intervenção e a entrada e saída de *peers* não altera seu comportamento e ainda são altamente tolerante à falhas por não apresentarem pontos únicos de falhas.

Dentre as iniciativas para preservação digital que utilizam redes P2P algumas se destacam como os Projetos LOCKSS, o BRICKS e o FARSITE.

O Projeto Lockss (Lots of Copies Keeps Stuff Safe), da Stanford University, que procura garantir a integridade das publicações eletrônicas pela manutenção de cópias em vários sites, checando periodicamente essas cópias para verificar a congruência informacional. Com a clonagem e o armazenamento distribuído. (REICH, 2009; ALTMAN, 2009)

O Projeto Farsite, desenvolvido pela Microsoft Research, visa construir um sistema de arquivos distribuídos de alta disponibilidade. Foi projetado para executar em grandes corporações ou universidades, numa rede de até 105 *peers* (topologia da rede pode ser ignorada) e ainda utiliza capacidade de processamento para proteger o conteúdo com criptografia.

O LOCKSS e o BRICKS são exemplos de sistemas de arquivamento confiável dos dados através de P2P. (BURNHILL, 2009) Ambos sistemas tratam seus repositórios com uma única probabilidade de falha, o que não modela exatamente a realidade. Além disso, esses sistemas não são compatíveis com o protocolo OAI-PMH. Sendo assim, a utilização de tais soluções não permite que exista integração entre Bibliotecas Digitais OAI.

Soluções baseados no paradigma P2P também estão sendo pesquisadas na Universidade Federal do Paraná, buscando suprir a necessidade de instrumentos que proporcionem a criação de Bibliotecas Digitais com baixo custo e com alto grau de confiabilidade e disponibilidade. Pois, as ferramentas existentes, como o LOCKSS e o BRICKS assumem que todos os objetos do sistema possuem a mesma importância, do ponto de vista de preservação. Tendo em vista que, em sua grande maioria, soluções de preservação digital são baseadas em replicação, é importante identificar os objetos que devem ser replicados com maior prioridade, ou seja, possuir mais réplicas no sistema. Assim, a UFPR propõe, um modelo de replicação confiável de conteúdo digital imutável para ser utilizado em sistemas de arquivamento a longo prazo, para ambientes distribuídos, também baseados em sistemas P2P. (VIGNATTI et al., 2009) Em seguida Seára (2009), propõe uma arquitetura que integra o modelo de replicação P2P ao protocolo OAI. O Objetivo final é manter a forma de disseminação de acesso ao conteúdo definida pelo protocolo OAI-PMH e criar um mecanismo, de baixo custo e alta confiabilidade, para preservação dos objetos. A

seção 3.1 apresenta o modelo de replicação desenvolvido na UFPR, a seção 3.2 descreve a integração na arquitetura OAI.

SISTEMA DE ARQUIVAMENTO DIGITAL SOBRE UMA REDE P2P ESTRUTURADA

Vignatti e colaboradores (2009) propõem um sistema de arquivamento digital totalmente distribuído aplicando o modelo e os algoritmos de criação de réplicas nas redes P2P. Neste sistema a comunicação dos nós (repositórios) é organizada em uma tabela hash distribuída (DHT) e múlltiplas funções hash são utilizadas para selecionar os repositórios que irão guardar as réplicas de cada item armazenado.

Inicialmente, Vignatti e colaboradores (2009) propuseram um modelo para repositórios de arquivamento que se caracteriza pela autonomia dos repositórios aos quais são associadas probabilidades independentes de falha. O modelo possui três componentes (Figura 3): uma rede de repositórios de arquivamento responsável pelo armazenamento das réplicas existe uma cooperação entre os repositórios e cada repositório possui uma capacidade de armazenamento limitada; itens ou objetos digitais, que é o material a ser preservado e por ultimo o componente denominado clientes, que são os usuários ou aplicações responsáveis por depositar e recuperar objetos digitais e também operar o sistema.

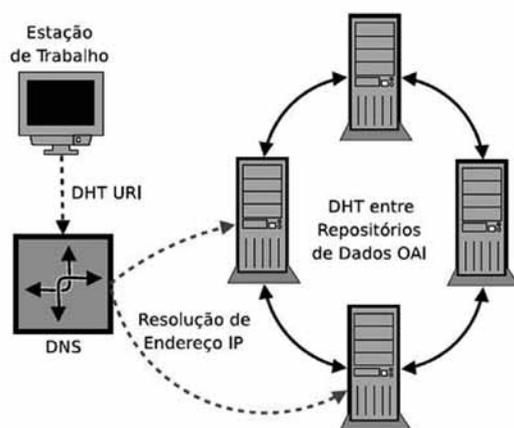



Figura 3 - Principais componentes do modelo de replicação
Fonte: Vignati e colaboradores (2009).

Para cada repositório existe um valor associado que indica a probabilidade de seus dados não serem perdidos em um determinado período de tempo, como mostrado pelas elipses da figura 3.

Após definir o modelo e seus componentes foram selecionados soluções algorítmicas para a replicação dos dados, baseados nesses estudos Vignati propôs também um Sistema de Arquivamento Digital implementado sobre a arquitetura de uma rede P2P.

Basicamente o sistema realiza as seguintes tarefas: (1) recebe requisitos de clientes para a inserção ou recuperação de objetos digitais no sistema. Na inserção de cada objeto, o cliente especifica uma confiabilidade que deseja que o objeto seja preservado; (2) dependendo da estratégia de replicação empregada e da confiabilidade de cada sítio, calcula o número de réplicas necessárias do objeto, e; (3) faz a auditoria de cada objeto digital comparando com as várias réplicas da rede, corrigindo-a se necessário e garantindo a preservação no longo prazo.

Na próxima seção será apresentado o estado da arte dos serviços oferecidos para arquivamento e acesso de informação para bibliotecas digitais, com o objetivo de traçar as perspectivas futuras de trabalho que estão sendo desenvolvidos pelo C3SL.

ARQUITETURA OAI PARA PRESERVAÇÃO DIGITAL

Seára (2009) propõe um modelo de arquitetura utilizando redes P2P (Peer-to-Peer) estruturadas, baseada no protocolo OAI. A proposta da arquitetura é manter as características atuais do protocolo OAI-PMH (Open Archives Initiative Protocol for Metadata Harvesting) e realizar, de forma transparente ao usuário, a preservação do conteúdo,considerando a importância do objeto para o administrador.

A popularidade do protocolo OAI-PMH entre as Bibliotecas Digitais mais recentes é baseada em sua simplicidade e arquitetura de baixo custo. Outro ponto que colabora, é a existência de *softwares* open source que o implementam, como DSpace e Eprints. Assim, a ideia central da arquitetura proposta por Seára (2009) é manter a forma de disseminação de acesso ao conteúdo definida pelo protocolo OAI-PMH e criar um mecanismo, de baixo custo, para preservação dos objetos.

A arquitetura baseia-se no pressuposto de que Bibliotecas Digitais OAI possuem o mesmo objetivo, portanto, podem colaborar para criação de um sistema que aumente a confiabilidade e disponibilidade de seus objetos.

Como mostrado na seção anterior, a Iniciativa Open Archives (OAI) desenvolve e promove padrões de interoperabilidade, com o objetivo de facilitar a disseminação eficiente de conteúdo. Uma federação OAI é claramente baseada na separação entre os Provedores de Dados (Data Providers – DP) e os Provedores de Serviço (Service Providers – SP).

A arquitetura proposta por Seára (2009) define a inserção dos Provedores de Dados OAI em uma Tabela Hash Distribuída (DHT) customizada que fornece funções de inserção, busca e auditoria dos objetos (Figura 4).

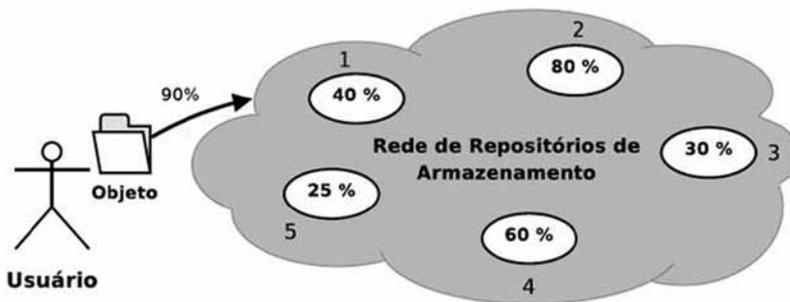



Figura 4 - Colaboração entre os repositórios de dados OAI através de uma DHT.
Fonte: Seára (2009).

Como pode ser observado, na figura 4, cada nó do sistema desempenha tanto o papel de Provedor de Dados como de DHT e, desta forma, realiza parte do trabalho de preservação e o gerenciamento de metadados. Ao receber a requisição do usuário, através de qualquer um de seus nós, a DHT realiza efetivamente a busca pelo objeto.

Por fim, o usuário acessa o objeto ou visualiza uma mensagem de erro, quando o objeto não for encontrado, através de uma página Web, gerenciada pelo DP, disponível no nó em que o conteúdo foi localizado.

De maneira geral, a principal contribuição da proposta de Seára (2009) foi a definição de uma arquitetura para preservação digital, com

o objetivo de aumentar a confiabilidade e disponibilidade dos objetos disponíveis em Bibliotecas Digitais.

A experiência do C3SL com arquiteturas de arquivamento, replicação e acesso digital para ambientes distribuídos, descritos nesta seção, comprova que os resultados positivos na combinação de tecnologias e iniciativas como OAI, redes e sistemas P2P tornam evidentes a necessidade de trabalhos futuros relacionados a serviços que podem ser disponibilizados a partir das infraestruturas apresentadas por Vignatti e outros (2009) e Seára (2009). Sendo assim, a seção a seguir descreve alguns serviços disponíveis na literatura para acesso e disponibilização de acervos para bibliotecas digitais.

O objetivo é contextualizar a necessidade de serviços para ambientes geograficamente distribuídos a partir da experiência obtida pelo C3SL, utilizando novas tecnologias e ferramentas de comprovada eficiência e baixo custo (desenvolvimento e recursos humanos), encontrados na literatura, para o desenvolvimento de arquiteturas de acesso ao acervo de bibliotecas digitais.

SERVIÇOS OFERECIDOS ÀS BIBLIOTECAS DIGITAIS

As bibliotecas digitais podem ser vistas como grandes repositórios de produção intelectual, sobretudo no que diz respeito à disseminação intelectual de grandes campos de pesquisas científicas.

A produção intelectual, quando relacionada a pesquisas científicas, possui algumas características que devem ser preservadas não apenas com o objetivo de serem disponibilizadas para a comunidade, mas também para servirem como base para novas pesquisas. Assim, as descobertas e acesso aos serviços oferecidos pelas bibliotecas digitais são o foco de muitas pesquisas em todo mundo.

São encontrados na literatura muitos estudos sobre serviços de acesso e recuperação de informação para bibliotecas digitais. Esses estudos compreendem as mais diversas técnicas e arquiteturas compu-

tacionais que vão desde serviços baseados em ontologias e semântica web até os serviços baseados em arquiteturas de sistemas distribuídos e serviços web.

Entre as propostas que envolvem ontologias e semântica web destaca-se o sistema PANIC. (HUNTER; CHOUDHURY, 2004) A Web Semântica tem como objetivo trazer significado aos dados armazenados nos servidores e páginas Web, tornando-os passíveis de serem processados e interpretados, por meio do uso de ontologias e regras de inferência.

O PANIC utiliza uma ontologia denominada Ontology Web Language for Service (OWL-S). A OWL-S permite a criação de ontologias e sua instanciação em sites da web, o objetivo é estabelecer uma estrutura para que as descrições sejam feitas e compartilhadas através de um serviço web. Os serviços web são programas de computador em rede que permitem processar e consumir informações, utilizando técnicas padronizadas para publicação, comunicação e acesso aos serviços.

O PANIC é um sistema desenvolvido para encontrar e invocar um serviço web de preservação digital, automaticamente, assim que detectar a obsolescência do objeto digital. A arquitetura do sistema é mostrada na figura 5. Analisando a figura é possível identificar quatro componentes principais do sistema: invocação, notificação, descoberta e provisão.

O componente de invocação é o conjunto de módulos de *software* localizado no servidor da empresa onde se encontra o serviço e fornece a interface de usuário para a tomada de decisões e definição de parâmetros associados a preservação.

O componente de notificação fornece uma interface de busca e recuperação da informação no registro de versão de *software*. Este componente é responsável por determinar a obsolescência potencial do objeto, registrando todas as informações sobre as ultimas autorias e versões.

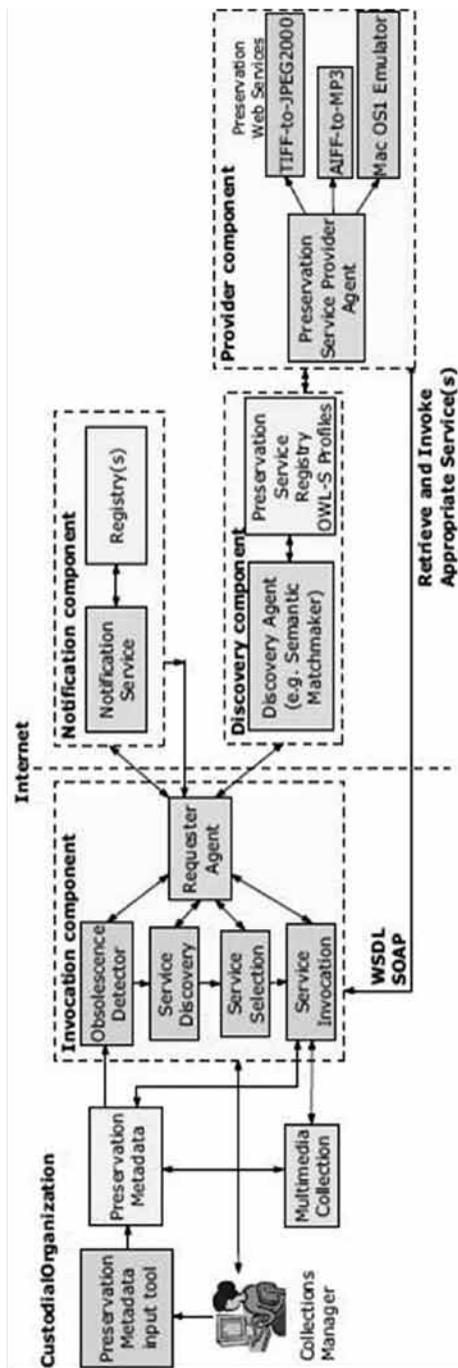



Figura 5 - Arquitetura do sistema PANIC
Fonte: Hunter e Choudhury (2004).

O componente de descoberta identifica o serviço do solicitante, procurando atender as suas especificações e necessidades. O componente envia as descrições do serviço de correspondência de volta ao solicitante e gera um documento eletrônico para apresentar os resultados da pesquisa.

Finalmente, o componente de provisão permite que o sistema chame automaticamente o melhor serviço web, depois de obter os resultados da pesquisa e as recomendações do solicitante.

A partir da integração de serviços com a semântica web, proposta por Hunter e Choudhury (2004), verifica-se que os provedores de serviços de preservação digital poderão se beneficiar dessa integração ao anotarem semanticamente seus serviços, com alguma linguagem de marcação semântica, como o OWL-S utilizada para o desenvolvimento do sistema PANIC. Essa integração possibilitou aos agentes de *software* descobrir, selecionar e executar Web Services de forma automática, com um grau reduzido de intervenção humana.

Outras propostas relevantes utilizando ontologias, semântica web e inferências podem ser encontradas na literatura. A proposta de Lin, Hong e Doerr (2008) destaca-se por apresentar uma plataforma de inferência robusta com a capacidade de inferir novos conhecimentos a partir de combinações de elementos distribuídos em diferentes fontes. Considerando que a maioria das pesquisas de recuperação de informação se concentra na coleta de documentos que correspondem ao mesmo conjunto de conceitos, é proposta uma solução para descobrir o conhecimento que não estão contidos em uma única fonte, mas sim em uma combinação fontes históricas.

Nesta mesma linha, Gonçalves, Fox e Watson (2008) propõem uma ontologia formal para que Bibliotecas Digitais definam os conceitos fundamentais, relações e regras axiomáticas que governam o domínio das Bibliotecas Digitais. Algumas aplicações práticas de ontologias são descritos, incluindo: a definição de uma taxonomia de serviços de biblioteca digital, a proposta de uma linguagem de modelagem de bibliotecas digitais, e as especificações de métricas de qualidade para avaliar Bibliotecas Digitais.

Em outro trabalho, Namba, Anzen e Okumura (2008) apresentam um

método que facilita obtenção de referência bibliográfica e a integração de busca de referências relacionadas a artigos, analisando as relações de citação entre eles. O modelo consiste em duas etapas. Na primeira etapa é realizada a detecção de frases contendo informação bibliográfica, e na segunda é feita a extração das informações bibliográficas nas sentenças. Os autores desenvolveram um sistema de recuperação de informação que forneceu dois métodos de recuperação de artigos de pesquisa e referências. Um método de recuperação por consulta, e outro por citação e as relações entre artigos e referências. Os resultados obtidos através de simulação comprovaram a eficiência da ferramenta.

Assim como o PANIC, outros trabalhos baseados em arquiteturas orientadas a serviços e serviços web também podem ser encontrados na literatura. Sorensen, Verelst e Sevaux (2008), propõem uma biblioteca baseada em serviços web, para um conjunto de informações para a comunidade cientifica de pesquisas operacionais.

Richardson, Srinivasan e Fox (2008) propuseram uma ferramenta denominada Knowledge Discovery System base on Web Services (KDWebS) baseada em serviços web, para facilitar a descoberta de conhecimento automatizado em uma biblioteca digital.

As mudanças rápidas em que as tecnologias de informação estão sujeitas têm reduzido consideravelmente a vida útil de objetos digitais. Neste contexto, Becker e outros (2008) apresentam uma arquitetura orientada a serviços e uma ferramenta de apoio à decisão que implementa a preservação do processo de planejamento para a preservação de objetos digitais. A proposta é integrar serviços de caracterização de conteúdo de um objeto digital com a ação de preservação.

Em todo mundo, iniciativas educacionais e governamentais mostram a preocupação e a necessidade de grandes bibliotecas e museus com a preservação do acervo digital. Mardis, Hoffman e Marshall (2008) realizaram um estudo em algumas bibliotecas digitais dos Estados Unidos da America (EUA), analisando o contexto atual e sugerindo soluções de melhorias baseadas em serviços web para acesso ao acervo digital. McMartin e outros (2008) realizaram um levantamento estatístico que demonstrou o uso mais frequente de bibliotecas digitais nas escolas de ensino médio dos EUA.

Na Europa, os estudos realizados Cousin, Chambers e Meullen (2008) descrevem as lições aprendidas na Biblioteca Europeia durante os dois últimos anos com a realização de padronização de dados e formatos para tornar viável a pesquisa de itens específicos nos acervos da Biblioteca Nacionais.

CONCLUSÕES E PERSPECTIVAS FUTURAS DA PRESERVAÇÃO DIGITAL

Serviços a serem fornecidos às bibliotecas digitais utilizando arquiteturas de AD para ambientes distribuídos.

Projetos de uma arquiteturas de *software* que centralize os acessos a uma base de dados distribuída por meio do acesso a uma biblioteca digital provê grandes benefícios para as pessoas beneficiadas com dados atuais e consistentes e aumento da confiabilidade e do aproveitamento das informações.

Considerando as arquiteturas de sistemas de AD para ambientes distribuídos e os serviços disponibilizados às bilbiotecas digitais apresentadas nas seções 3 e 4 é possível fazer projeções futuras sobre possíveis arquiteturas de acesso orientados a serviços (serviços web) e sistemas P2P para ambientes distribuídos. Pois, a disponibilização de serviços unificados e centralizados para acesso à uma biblioteca digital utilizando-se uma arquitetura de um serviço web e sistemas P2P possibilitará que os usuários destes serviços façam isso de forma independente da plataforma ou da linguagem de programação.

O objetivo é fornecer uma arquitetura para aumentar o reaproveitamento de informações confiáveis, baseados em uma solução semi-pronta baseada em serviços web.

As arquiteturas orientadas a serviços e serviços web apresentam-se como uma forma de interconexão de aplicações entre sistemas computacionais através da web. Além disso, por possuírem uma arquitetura eminentemente aberta e padronizada, os serviços web possuem um grande potencial de uso para a computação distribuída.

A possibilidade de tecnologias relacionadas a SOA e serviços web

ser apenas um modismo ou se realmente será algo que se tornará uma realidade comercial já foi superada, conforme Vaughan-Nichol (2002), pois praticamente todas as soluções apresentadas no mercado e que possuem como base a Web, são completamente ou em partes, desenvolvidos utilizando-se alguma arquitetura de desenvolvimento que envolve serviços web.

Diante das possibilidades e tecnologias apresentadas nas seções anteriores relacionados à preservação digital, tecnologias e arquiteturas de arquivamento digital e serviços disponibilizados torna-se relevante estudos que possam utilizar, de forma eficiente e eficaz, estes recursos.

Desta forma, vários trabalhos começam a ser desenvolvidos pelo C3SL para aprofundar os conceitos da preservação digital e das bibliotecas digitais e identificar soluções relevantes que contribuam para a integraçãode arquiteturas de arquivamento digital para ambientes distribuídos utilizando redes P2P e SOA´s, como os serviços web e os modelos proveniência de dados.

REFERÊNCIAS

ALTMAN, J. et al. Taking Charge: preserving our digital heritage, Part II. **Against the Grain**, v. 21, n. 2, April 2009. Disponível em: <http://www.against-the-grain.com/d/TOCIssue?&volsearch=21&issuesearch=2>. Acesso em: 20 July 2010.

BECKER, C. et al. Plato: a service-oriented decision support system for preservation planning. In: JOINT CONFERENCE ON DIGITAL LIBRARIES, 8., 2008. Pittsburgh, Pennsylvania. **Proceedings**... Pittsburgh, ACM/IEEE-CS, 2008.

BURNHILL, P. et al. Taking Charge: preserving our digital heritage, Part I. **Against the Grain**, v. 21, n. 1, Feb. 2009. Disponível em: <http://www.against-the-grain.com/d/TOCIssue?&volsearch=21&issuesearch=1.

CCSDS - Consultative Committee for Space Data Systems. **Reference model for an Open Archival Information System (OAIS) - Blue Book**. Washington: National Aeronautics and Space Administration, 2002. Disponível em: <http://public.ccsds.org/publications/archive/650x0b1.pdf>. Acesso em: 17 Sep. 2009.

COUSINS, J.; CHAMBERS, S.; MEULEN, E. Uncovering cultural heritage through collaboration. **International Journal on Digital Libraries**, v. 9, n. 2, p. 125-138, Nov. 2008.

FERREIRA, M. Automatic evaluation of migration quality in distributed networks of converters. **Bulletin of the IEEE Technical Committee on Digital Libraries** (TCDL), v. 2, n. 1, 2006.

GONÇALVES, M.; FOX, E.; WATSON, L. Towards a digital library theory: a formal digital library ontology. **International Journal on Digital Libraries**, v. 8, n. 2, p. 91-114, April. 2008.

HUNTER, J.; CHOUDHURY, S. A semi-automated digital preservation system based on semantic web services. In: JOINT CONFERENCE ON DIGITAL LIBRARIES, 4., 2004. Tucson, Arizona. **Proceedings...** Tucson: ACM/IEEE-CS, 2004.

LAVOIE, B. F. **The open archival information system reference model:** introductory guide. Dublin, USA: OCLC, 2004. Technology Watch Report Watch Series Report 04-01.

LEE, K.; SLATERY, R. L.; McCARY, Y. The state of the art and pratice in digital preservation. **Journal of Research of the National Institute of Standards na Technology,** v. 107, n. 1, p. 93-106, 2002.

LIN, C.; HONG, J.; DOERR, M. Issues in an inference platform for generating deductive knowledge: a case study in cultural heritage digital libraries using the CIDOC CRM. **Int. J. on Digital Libraries**, v. 8, n. 2, p. 115-132, 2008.

LORIE, R. A methology and system for preserving digital data. In: JOINT CONFERENCE ON DIGITAL LIBRARIES, 2., 2004 Portland, Oregon, USA. **Proceedings...** Portland, ACM/IEEE-CS, 2004.

MARDIS, A. M.; HOFFMAN, E. S.; MARSHALL, T. E. A new framework for understanding educational digital library use: re-examining digital divides in U.S. schools. **International Journal on Digital Libraries**, v. 9, n. 1, p. 19-27, 2008.

McMARTIN, F. et al. The use of online digital resources and educational digital libraries in higher education. **Journal International Journal on Digital Libraries**. v. 9, n. 1, p. 65-79, Aug. 2008.

NAMBA, H.; ANZEN, N.; OKUMURA, M. Automatic extraction of citation information in Japanese patent applications. **International Journal on Digital Libraries**.v. 9, n. 2, p. 151-161, Nov. 2008.

REICH, V. Distributed Digital Preservation. In: INDO-US WORKSHOP ON INTERNATIONAL TRENDS IN DIGITAL PRESERVATION, 2009. Pune, India. **Proceedings...** Pune, 2009.

RICHARDSON, W.; SRINIVASAN, V.; FOX, E. Knowledge discovery in digital libraries of electronic theses and dissertations: an NDLTD case study. **International Journal on Digital Libraries**, v. 9, n. 2, p. 163-171, Nov. 2008.

ROTHENBERG, J. **Avoiding technological quicksand: finding a viable technical foundation for digital preservation: a report to the Council on Library and Information Resources**. Washington, DC: Commission

on Preservation and Access and Council on Library and Information Resources,1999.

SEÁRA, E. **Uma arquitetura OAI para a preservação digital utilizando redes peer-to-peer estruturadas**. 2009. 90 f. Dissertação (Mestrado em Ciência da Computação) – Departamento de Informática, Universidade Federal do Paraná, Curitiba, 2009.

SORENSEN, K.; VERELST, J.; SEVAUX, M. Stimulating information sharing, collaboration and learning in operations research with labor. **International Journal on Digital Libraries**, v. 8, n. 2. p. 79-90, Apr. 2008.

THIBODEAU, K. Overview of technological approaches to digital preservation and challenges in coming years. Washington D.C., 2002. Disponível em: < http://www.imaginar.org/dppd/DPPD/40%20pp%20Overview%20of%20 Technological%20Approaches.pdf>. Acesso em: 21 jul. 2010.

VAUGHAN-NICHOL, S. J. Developing the distributed-computing OS. **IEEE Computer**, v. 35, n. 1, p. 19-21, Jan. 2002.

VIGNATTI, T. et al. Long-term Digital Archiving Based on Selection of Repositories Over P2P Networks. INTERNATIONAL CONFERENCE ON PEER-TO-PEER COMPUTING, 9., 2009. Seattle, Washington, USA. **Proceedings…** Seatle: IEEE Communications Society, 2009.

Porta curtas e curta na escola – repositórios

modelo para difusão da produção cultural brasileira, ao público e para escolas, via internet

Vanessa Souza

INTRODUÇÃO

Alcançar o grande público já é um desafio para os longas-metragens brasileiros, que dirá para os filmes de curta-metragem, geralmente renegados a uma categoria menor dentro de festivais de cinema. Mas o Porta Curtas, patrocinado pela Petrobras via Lei Rouanet de Incentivo à Cultura está revertendo essa situação e, desde sua estreia, em agosto de 2002, já promoveu perto de doze milhões de exibições de curtas na internet para espectadores de mais de 2.600 cidades em 154 países, o que equivale a 100 mil sessões em salas de cinema completamente lotadas.

O repositório Porta Curtas[1] surgiu com a missão de catalogar para pesquisas na web todo o acervo brasileiro de filmes de curta-metragem e exibir gratuitamente uma ótima e farta seleção. Atualmente, o serviço disponibiliza mais de 700 curtas para assistir *online*, além de um universo catalogado de 6 mil fichas-filme e 25 mil profissionais do cinema.

No Porta Curtas, os filmes estão em constante exibição e encontram público onde um curta-metragem sequer sonhou em chegar – locais onde simplesmente não existem festivais de curtas-metragens.

Entre as grandes metas do projeto, destacam-se: oferecer ao internauta interessado em curtas a possibilidade de pesquisar, achar e assistir os filmes antes só (e dificilmente) acessíveis nos festivais; e exibir curtas-metragens para o público comum, os "não-cinéfilos", surpreendendo-os com excelentes filmes brasileiros, democratizando o gênero, criando um canal de exibição para um segmento desprestigiado pelo mercado exibidor e, ao mesmo tempo, formando novas plateias para o cinema nacional.

Por ser um serviço de difusão de informações, o Porta Curtas propõe a reversão do conceito de mídia e exclusividade e estabelece parcerias com editores de outros portais, sites profissionais, *blogs*, jornais e revistas *online*, que são estimulados a promover e divulgar curtas-metragens em sua programação. Esta iniciativa tem tanto sucesso que vários parceiros criaram sessões permanentes para exibição de curtas-metragens em seus canais web.

O acervo do Porta Curtas transcende o entretenimento e é também utilizado como conteúdo em salas de aula. O Curta Na Escola, extensão pedagógica do projeto, estreou em abril de 2007 com a finalidade de constituir uma Rede Nacional de Aprendizagem em torno da construção colaborativa de conteúdos relacionados ao uso dos curtas-metragens brasileiros em escolas de todo o país. Para isso, o projeto apresentou duas inovações: o desenvolvimento do site www.curtanaescola.org.br e a Coleção DVD Curta na Escola. Voltaremos ao Curta Na Escola mais adiante.

1 www.portacurtas.org.br

Os números de audiência e as críticas positivas recebidas ao longo dos anos de atividade do Porta Curtas e também de sua extensão pedagógica, são extremamente positivos e comprovam a consolidação do projeto como a mais importante referência em difusão de curtas-metragens, no Brasil ou no exterior, agora com benefícios no setor educacional brasileiro.

A seguir, um recorte dos muitos elogios recebidos durante os anos e atividade do Porta Curtas :

> Patrocinado pela Petrobrás, o site Porta Curtas é uma das melhores novidades da web nos últimos tempos. São 25 curtas-metragens da ótima safra atual do cinema brasileiro, disponibilizados para *download* em seu formato original e na íntegra. (Marina Monzillo, em *Isto É Gente* de 09/09/2002)

> Em vez de sair zapeando a TV por assinatura, alguns milhares de cinéfilos plugados estão correndo atrás do prejuízo de não verem curtas-metragens no cinema no www.portacurtas.com.br... Esta semana, o site chegou à marca de 405 mil curtas exibidos na Internet. (Eduardo Simões, *O Globo*, Segundo Caderno de 6/06/2003)

> Gostaria de cumprimentá-los pelo Porta Curtas . É uma iniciativa fascinante de divulgar nossa cultura e nossa arte. (Prof. Roberto Grobman, usuário do serviço - 2004)

> Recebi ótimos retornos de pessoas que descobriram e assistiram ao filme pelo site. Quanto ao carinho pelo site, esse é incondicional pela importância, seriedade e boa vontade do trabalho de vocês. (Aurélio Aragão, diretor de "Quando um Burro Fala..." - 2005)

> Curto muito o site de vocês e espero que o Portas Curtas não saia do ar nunca! Obrigado por nos mostrar que, mesmo sendo uma criança grande, é sempre bom saber e ficar por dentro da

nossa cultura que tão poucas pessoas conhecem! (Rafael Luis, usuário do serviço - 2006)

Quero parabenizá-los pela bela iniciativa e projeto. Nós professores precisamos deste incentivo e mais que isto, destas ferramentas, destes exemplos práticos de utilização das mídias, já que muito se fala em tecnologia educacional, de modernização, mas pouco se faz em termos práticos e objetivos que auxiliem ao professor na sua prática docente diária. (Prof. Claudia Dutra Gallo, Colégio Estadual Poeta Mário Quintana - Nilópolis/RJ – 2007)

Conhecemos o belo trabalho do Porta Curtas e também o Curta na Escola, ações que merecem o maior respeito e incentivo de nossa parte. (Prof. João Luís Almeida Machado, Editor do site Planeta Educação - 2008)

Sou um grande admirador da produção nacional, sobretudo do formato curta-metragem. Inicialmente gostaria de parabenizá-los por este excelente canal de divulgação que faço questão de divulgar a todos que conheço. (Valdir Villela Jr., usuário do serviço, 2009)

BREVE HISTÓRICO

O falecimento do goleiro Barbosa, personagem central da histórica derrota brasileira para o Uruguai na Copa de 1950, foi notícia em grandes jornais impressos e também nos principais sites jornalísticos da Internet brasileira, em abril de 2000.

Nessa ocasião, a empresa Synapse Produções, provedora de conteúdo audiovisual para o mercado brasileiro e internacional, fez um acerto com o Provedor de conteúdo UOL para a publicação do curta-metragem "Barbosa", estrelado por Antonio Fagundes e dirigido por

Ana Luíza Azevedo e Jorge Furtado, em seu portal de esportes, logo abaixo do lead da matéria.

O resultado impressionante – 27.654 exibições contabilizadas – comprovou o potencial da Internet para difusão do curta-metragem brasileiro e, detalhe importante, a vantagem de se promover a exibição de forma segmentada, ou seja, o filme ir de encontro ao seu espectador, e não o contrário.

A partir dessa experiência bem-sucedida, e do desejo do jornalista Julio Worcman de organizar um banco de dados catalogando a produção brasileira de curtas-metragens produzidos a partir da década de 1980 – primavera do curta-metragem nacional, o projeto Porta Curtas surgiu com a pretensão de oferecer aos editores de internet conteúdo rico que pudesse ser associar aos assuntos de suas páginas.

O interesse da empresa em patrocinar o projeto (incentivo via Art 18° da Lei Rouanet) foi fator determinante para o início das atividades do Porta Curtas, que incorporou a marca do patrocinador ao seu nome. Nasceu então o serviço Porta Curtas , em agosto de 2002, quase três anos antes do Youtube – fevereiro de 2005.

Para marcar sua estreia com uma celebração ao curta-metragem nacional, o Porta Curtas promoveu votação para eleger os dez mais importantes filmes brasileiros de curta-metragem de todos os tempos. Dez pessoas ligadas ao cinema, entre críticos, realizadores e produtores, escolheram os 10+, que passaram a integrar o acervo permanente do site.

Os cinco curtas-metragens mais votados foram projetados nos eventos de lançamento do serviço Porta Curtas: o primeiro realizado durante o Festival Internacional de curtas-metragens de São Paulo, no Museu da Imagem e do Som, no dia 26 de agosto; e o segundo no Cinema Odeon, no Rio de Janeiro, dia 9 de setembro. O público estimado participando dos eventos, seguidos de coquetel, foi de mais de 500 pessoas do setor audiovisual e formadores de opinião, jornalistas, etc.

A partir de então, o Porta Curtas não parou de crescer. Seja devido ao acervo riquíssimo e em constante atualização; pela organização do banco de dados; por acompanhar as tendências e evoluções tecnológicas; ou por suas ferramentas interativas que estimulam o consumo

dos curtas-metragens, o projeto alcançou expressivo espaço na web, consolidando-se ao mesmo tempo com canal de entretenimento e veículo de difusão cultural.

UNIVERSO CATALOGADO E ATIVADADE DOS USUÁRIOS

Com a incrível média mensal de 150 mil visitas, 2 milhões de pageviews e 200 mil exibições de curtas, o Porta Curtas elevou o curta-metragem brasileiro, antes restrito ao circuito de festivais presenciais, ao *status* de produto cultural, superando em retorno de audiência a grande maioria dos longas-metragens brasileiros da produção recente.

Os dados demonstrando a envergadura do serviço de difusão cultural realizado pelo Porta Curtas , inclusive com detalhes sobre o intenso uso do serviço pelo público (pesquisas realizadas, uso das ferramentas interativas, elogios de usuários etc.), podem ser facilmente acompanhados através de dois relatórios disponíveis *online* e abertos ao público:

- Estatísticas *online* - http://portacurtas.com.br/estatisticas1.asp (dados atualizados em tempo real)
- Relatórios ao Patrocinador - http://portacurtas.com.br/relatorios/2009_2/

Eis um pequeno apanhado de dados (até dezembro de 2009) sobre o universo catalogado e a atividade dos usuários no serviço Porta Curtas :

UNIVERSO CATALOGADO

- Curtas Para Pesquisar: 5861
- Curtas Para Assistir: 730
  - Animação - 114
  - Conteúdo Adulto - 52
  - Documentário - 212
  - Experimental - 68
  - Ficção - 358
  - Grande Prêmio Vivo – 16

- Artistas e Técnicos - 24.157
  - Atores – 8.641
  - Diretores – 4.706
  - Fotografia - 3.489
  - ANIMADORES - 442

ATIVIDADES DOS USUÁRIOS

- Usuários Cadastrados: 187.679
- Assinantes informativo Curta Clube: 112.508
- Coexibidores sites profissionais cadastrados: 416
- Coexibidores páginas pessoais cadastrados: 8.114
- Exibições de curtas até o momento: 11.672.930
  - Exibições a partir do Porta Curtas: 7.237.217
  - Exibições por outros sites profissionais: 4.027.161
  - Exibições por outros sites pessoais: 408.553
- Envio de curtas por usuários via e-mail (torpedos): 241.701
- Comentários aos filmes postados por usuários: 7.226
- Votos (rating) de usuários aos curtas: 234.271
- Pastas de cinemateca com curtas arquivados: 92.553
- Número de usuários com curta(s) arquivado(s) em cinemateca: 79.862
- Curtas arquivados em pastas de cinemateca: 353.223
- Fichas do acervo com alguma interação por usuários: 5.150
- Pesquisas respondidas: 26.817.845
- Por obra completa de um nome em ficha técnica: 23.969.599
- Buscas genéricas respondidas: 2.520.618
- Buscas detalhadas respondidas: 128.445
- Buscas respondidas na janela de exibição: 199.183
- Visualizações de fichas completas de curtas: 2.295.599
- *Downloads* de roteiros para leitura: 66.930

Um dado relevante a se destacar é que quase 90% das fichas catalogadas registraram alguma interação por parte dos usuários desde a

inauguração do serviço. Considerando que a ficha de um filme sofre interação quando é arquivada em alguma Cinemateca, ou recebe comentário, ou voto, ou foram recomendadas por email ou o filme foi assistido, esse dado é, sem dúvida, surpreendente, especialmente em se tratando de um setor ainda pouco difundido como o dos curtas-metragens.

O Porta Curtas prova, desta maneira, que uma catalografia bem realizada, amiga do usuário e acessível pela Internet faz com que as informações sobre as obras culturais circulem e sejam utilizadas pelos cidadãos.

USABILIDADE DO SERVIÇO

Quando o Porta Curtas estreou, em 2002, praticamente não havia oferta de vídeos na web. Além disso, a qualidade e o tamanho da janela de exibição eram bem inferiores ao que estamos acostumados hoje. Nos últimos anos, porém, com a popularização do acesso à banda larga e o fenômeno Youtube, houve a consolidação dos vídeos na web. Prova disso é que os principais provedores de conteúdo lançaram seus canais de vídeo – Globo Vídeos, TV UOL, IG TV, Terra TV.

Mas, apesar da oferta crescente de vídeos gratuitos na internet, o Porta Curtas mantém seu potencial de difusão da filmografia de curtas-metragens nacionais. O grande sucesso é atribuído a quatro pilares fundamentais de sustentação:

- Web design e conteúdo editorial
- Ferramentas interativas
- Resultados de busca ricos (por consequência da inclusão de meta dados *rich media*)
- Sistema automatizado de parcerias de difusão

WEB DESIGN E CONTEÚDO EDITORIAL

O site Porta Curtas possui design rico e atraente. Utiliza cores fortes e vibrantes, ilustrações criativas e ferramentas que estimulam a

interação, o serviço oferece uma interface ao mesmo tempo agradável e inquietante. Inquietante, pois ao acessar a página o usuário fica diante de uma infinidade de opções: pode assistir um curta, comentar, pesquisar festivais, ler notícias na sessão Curtinhas, navegar em links interessantes relacionados ao mundo do cinema etc. Essa infinidade de informações e opções poderia ser um problema, com tantas possibilidades é fácil se perder ou perder o foco. No entanto, a arquitetura do site é simples, funcional e bem organizada.

Navegando no Porta Curtas, o usuário está sempre a um clique de seleções de curtas-metragens por tema, por mais vistos, mais cotados e recomendados por outros usuários. Pode, ainda, acessar a ficha completa das obras, contendo gênero, ano de produção, cor, bitola, país e outras informações relevantes.

O acervo do Porta Curtas, além de riquíssimo – mais de 700 curtas representando o melhor da produção nacional do gênero, está

em constante crescimento. Toda semana um curta novo estreia na programação.

O respeito ao autor é uma das premissas do projeto, por isso, apesar de gratuita, a exibição dos filmes é sempre autorizada através de contrato de licenciamento, onde os detentores dos direitos autorais, em geral, são remunerados pela obra.

O contrato de licenciamento oneroso dos filmes é celebrado por prazo indeterminado, porém, após dois anos, o autor pode solicitar o cancelamento da exibição. Um indício do sucesso do projeto entre os realizadores é que a quase totalidade dos curtas contratados permanece em exibição após esse período.

A critério do autor, cada contrato pode, ainda, conter cláusula de comercialização não exclusiva, que abrange a venda ou aluguel do filme para terceiros através do Porta Curtas. Os idealizadores do projeto acreditam que a falta de exclusividade não prejudica o serviço, pelo contrário, seu objetivo primordial é a democratização do conteúdo, por isso, não cabe limitar seu alcance pelo público.

Os visitantes do Porta Curtas também podem ser cadastrar para receber informativos semanais, denominados Curta Clube, com indicações de curtas para assistir, dicas e notícias referentes a festivais de cinema e filmes premiados. O sistema de cadastramento permite aos usuários indicarem suas preferências em atores, gêneros ou assuntos, garantindo o recebimento de informações de seu interesse específico e chamando-os assim a acessar o portal.

FERRAMENTAS INTERATIVAS

O Porta Curtas incentiva o consumo interativo dos curtas do acervo através de ferramentas específicas para atrair o espectador:

– Janela de exibições de curtas

Os filmes de curta-metragem disponibilizados no Porta Curtas podem ser assistidos nos formatos Flash Vídeo (progressive *download*),

Windows Media Vídeo (*Streaming*) e QuickTime (formato padrão do iPhone), de acordo com a preferência do usuário. O usuário pode optar por assistir os curtas também em tela cheia.

– Dê a sua nota

Os usuários podem cotar os filmes assistidos dando de uma a cinco estrelas. A votação média serve de referência para novos usuários.

– Minha Cinemateca (pré-seleção de curtas)

Todos os curtas do acervo do site podem ser agrupados em pastas, denominadas Cinematecas. O sistema também permite que os usuários enviem suas cinematecas por email, multiplicando o número de fãs e conhecedores de curtas-metragens.

Eis um exemplo de Cinemateca de usuário, contendo pastas nomeadas pelo próprio frequentador do Porta Curtas .

O Porta Curtas conta com 93 mil pastas de cinemateca, somando 350 mil curtas arquivados. Dos 180 mil usuários cadastrados no serviço, 80 mil utilizam a ferramenta para arquivar seus curtas preferidos.

– Torpedos (envio de recomendações de filmes por e-mail)

Os torpedos podem ser enviados a partir da ficha de um filme, ou através das Cinematecas. As indicações de curtas por email podem seguir para um ou mais destinatários simultaneamente.

Desde o início das operações do serviço, 241.701 recomendações de curtas foram enviadas via email, gerando também retorno à exibição dos filmes.

– Comentários de usuários aos filmes

Todo comentário de usuário passa por uma pré-aprovação da equipe editorial. Quando liberado, tem uma cópia enviada para o diretor, e

outra publicada no site associada ao filme em questão.

A imagem a seguir mostra alguns comentários postados por usuários ao curta *Ilha das Flores*, de Jorge Furtado.

Assim como as votações, os comentários servem de referência para outros usuários na seleção dos filmes, inclusive abrindo a possibilidade de criação de uma espécie de fórum, através do mecanismo que permite que novos comentários sejam acrescentados a outros.

– Curta Clube (informativo semanal)

Os usuários cadastrados no Porta Curtas têm a alternativa de re-

ceber um informativo semanal sobre estreias e outras novidades do site e do mundo dos curtas-metragens. O serviço já conta com mais de 110 mil assinantes do informativo Curta Clube.

– *Download* de Roteiros de curtas

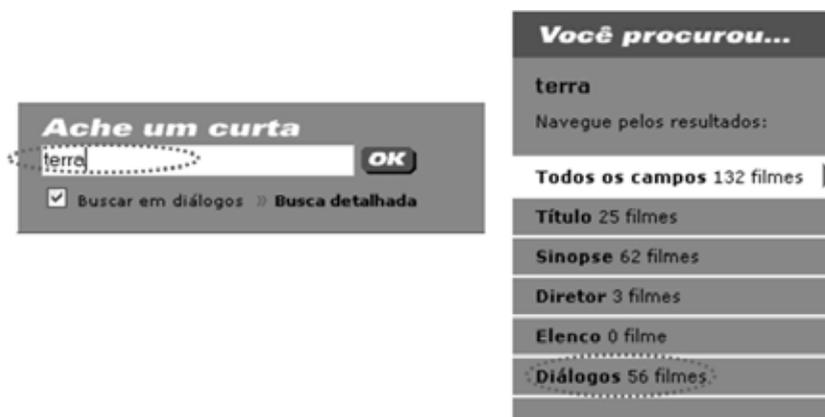
Os usuários do Porta Curtas podem baixar, para ler em seus computadores ou imprimir, a íntegra dos roteiros dos curtas disponibilizados para exibição no projeto. A ferramenta tem grande uso para estudantes de cinema e aficionados.

– Roleta

Ferramenta divertida, oferece um curta-metragem de maneira aleatória aos usuários.

RESULTADOS DE BUSCA RICOS

O banco de dados do Porta Curtas serve como um censo dos diretores, atores e técnicos que já estiveram ou estão envolvidos na atividade de produção de curtas metragens. Para facilitar a navegação entre eles, todos os profissionais catalogados são hiperlinks que geram listas com os curtas-metragens nos quais cada nome participou, tornando o sistema de catalografia convidativo à pesquisa e muito amigo do usuário.

Além disso, através da ferramenta "Ache um Curta", que tem localização privilegiada no topo de todas as páginas do site, o usuário pode fazer sua pesquisa, digitando palavras como assunto, título da obra, ator, diretor, diretor, técnicos, sinopse dos filmes, roteiro e diálogos.

A catalografia é uma das bases do Porta Curtas, por isso o sistema apresenta constantes atualizações para oferecer ao usuário opções de buscas cada vez mais completas, como a busca por palavras nos diálogos e a busca amigável por filtros.

Desde o início das operações o serviço Porta Curtas já registrou 27 milhões de visualizações de páginas de resultados de pesquisas (buscas) genéricas. O fato revela a plena usabilidade do serviço, com usuários

pesquisando curtas por títulos, por nomes de atores, de diretores ou técnicos envolvidos nas produções, ou ainda por gênero, tema, palavras ou frases nas sinopses ou nos roteiros completos dos curtas.

SISTEMA AUTOMATIZADO DE PARCERIAS DE DIFUSÃO

A ideia das parcerias de difusão do Porta Curtas consiste em oferecer curtas cujo tema e/ou elenco tenham afinidade com o conteúdo editorial de portais de conteúdo, *sites*, *blogs*, jornais e revistas *online*.

Em uma chamada na página principal do Porta Curtas, editores são estimulados a se cadastrar como "parceiros difusores" para publicar gratuitamente qualquer conteúdo do Porta Curtas.

Parceiros de portais ou sites de grande visitação de público são categorizados como parceiros profissionais e coassinam as exibições a partir de suas páginas (ver exemplo abaixo), aumentando a empatia entre os editores *online* e o projeto, e proporcionando aos usuários desses portais um acesso direto ao conteúdo do serviço.

Já são mais de 416 parceiros profissionais que mantém diálogo permanente com os editores do Porta Curtas para indicação de curtas cuja temática esteja de acordo com os conteúdos em pauta nos veículos.

Esta iniciativa tem tanto sucesso que vários parceiros profissionais criaram canais temáticos, oferecendo seleções especiais do acervo Porta Curtas . Dentre estes, podemos destacar *UOL*, *IG*, *O Dia*, *O Globo*, *Folha de São Paulo*, *Caros Amigos*, entre outros.

Confira a seguir a coluna Porta Curtas no UOL Cinema:

O uso regular do acervo do Porta Curtas em grandes portais e sites profissionais propiciam uma divulgação jamais experimentada por esse segmento da produção cultural brasileira.

A variedade de temas do acervo proporciona o alcance de um grande público interessado em exibir curtas dos mais diferentes assuntos, seja como uma opção de entretenimento ou ilustrando textos.

Dessa forma, milhares de editores de *blogs* e páginas pessoais também incorporam aos seus canais os conteúdos disponibilizados pelo Porta Curtas , contribuindo, assim, para a difusão capilar do curta metragem brasileiro. Atualmente, já são mais de 8 mil *blogs* cadastrados como parceiros difusores do Porta Curtas .

Para reportar a audiência dos curtas nos sites parceiros, o Porta Curtas disponibiliza dois relatórios:

- Exibições por exibidor: lista todos os curtas exibidos no site parceiro em determinado período; e informa a audiência total do site e a de cada curta.
- Exibições por filme: informa a audiência do curta selecionado, dentro de um período delimitado pelo parceiro difusor.

Os relatórios estão disponíveis na área administrativa do Porta Curtas , e o acesso exige login e senha de cadastro.

Uma das novidades recentes oferecida aos parceiros difusores é o Player Embed. Essa ferramenta permite a publicação de um tocador de vídeo direto na página cadastrada pelo parceiro, sem *link* para a janela de exibições do Porta Curtas . O modelo de publicação a partir de um código HTML adicionado nas páginas é o mesmo apresentado pelo Youtube, só que, neste caso específico, com mais funcionalidades.

No Player Embed do Porta Curtas é possível, além de assistir o filme, acessar informações completas sobre a obra, comentar e ler comentários, indicar por email, buscar por outro filme do acervo, fazer o *download* do roteiro, visualizar aplicabilidades pedagógicas e relatos de experiência com o uso do filme em salas de aula, entre outras coisas.

A ilustração a seguir é um recorte da publicação do curta-metragem *Eu sou um Panda*, de Gurcius Gewdner, no parceiro difusor Filmescópio.

MELHORIAS E INOVAÇÕES

O Porta Curtas surgiu como um dos mais inovadores portais de exibição de conteúdo audiovisual na internet, apresentando conteúdo e ferramentas na época pouco ou nada difundidas. Anos depois, a exibição de vídeos na web é mania em todo o mundo, facilitada inclusive pelo aumento constante do número de computadores com conexão em banda larga.

Para atender a esse novo contexto virtual e aos seus usuários, muito mais exigentes, um novo sistema encontra-se em desenvolvimento. Esse sistema está sendo pensado a partir da constante observação das necessidades dos usuários aliada a uma extensa pesquisa nos novos mecanismos e hábitos de usuários surgidos com a web 2.0.

O novo sistema Porta Curtas trará evoluções das ferramentas atuais, será internacionalizado (com uma primeira versão em Inglês) e deverá englobar um sistema de e-commerce automatizado para venda de curtas-metragens.

Apesar da modernização do site ser de extrema importância, um passo ainda mais importante rumo à democratização do acesso à produção nacional de curtas-metragens de qualidade foi dado com o lançamento do projeto Curta Na Escola, em 2006.

CURTA NA ESCOLA

O projeto Curta na Escola nasceu com os entraves típicos da exclusão digital na era da informação, agravados pela dimensão de um país que conta com áreas de isolamento cultural e tecnológico. A questão era: como conectar pontos tão isolados a um centro comum de informação?

A implantação de uma inovadora estratégia de difusão e de participação colaborativa de educadores fez com que em um só lance dois grandes objetivos fossem atingidos:

- Propagar a arte e a cultura brasileira em lugares aonde não há salas de exibição cinematográfica;
- Propiciar a inclusão digital participativa de professores e alunos da rede pública de ensino.

Em pouco mais de 2 anos de atividade, o projeto já reúne 262 curtas com aplicabilidades pedagógicas, mais de 700 relatos (colaborações estruturadas) de professores sobre suas experiências em salas de aula e 296 pareceres indicando aplicabilidades pedagógicas de filmes.

O curta-metragem brasileiro apresenta diferenciais educacionais importantes, quando comparado a outros conteúdos:

- Este tipo de conteúdo, por ser produzido no Brasil e por realizadores brasileiros, representa nossa sociedade e a nossa cultura;
- A qualidade da produção brasileira de filmes neste formato é reconhecida em todo o mundo por sua excelência;
- Sua curta duração, geralmente próxima de 15 minutos, faz com que os conteúdos sejam ideais para utilização em sala de aula, permitindo que os filmes sejam utilizados como "porta de entrada" de um assunto, fonte adicional de informação, pretexto para debater um tema ou para "coroar" o final de um projeto.

Sendo assim, a criação de um serviço de Internet destinado à construção colaborativa de conteúdos pedagógicos em torno do uso de curtas-metragens em salas de aula é fundamental. Com essa intenção, nasceu o projeto Curta na Escola e seu desdobramento, a Coleção Curta na Escola.

WWW.CURTANAESCOLA.ORG.BR

DVD COLEÇÃO CURTA NA ESCOLA

A missão do Curta na Escola é incentivar a produção de conhecimento cooperativamente e, para isso, lançou mãos de processos pioneiros, propiciando elevar o professor à condição de autor de conteúdos conexos aos filmes apresentados.

No site Curta na Escola, educadores cadastrados compartilham suas vivências em torno da utilização dos curtas em sala de aula através de comentários aos filmes, discussões no fórum e, principalmente,

do envio de relatos de suas experiências com a exibição de filmes aos alunos. Tais relatos integram um grande banco de experiências educacionais, permanentemente aberto para consultas públicas.

DETALHE DE UM RELATO

BANCO DE RELATOS

Para enriquecer a qualidade de exibição dos curtas em salas de aula (alternativamente à exibição via Internet), a equipe pedagógica do Curta na Escola selecionou 16 filmes com alto potencial didático, todos apropriados para uso em todos os níveis de ensino, para compor os DVDs Coleção Curta na Escola - Vol. 1 e 2.

Cada filme incluído na Coleção possui pareceres pedagógicos elaborados por pedagogos especializados – tais como Manoel Morán, Beatriz Rizek, entre outros – sugerindo roteiros passo-a-passo para sua utilização com os alunos.

Sabe-se que o Ministério da Educação, através do Projeto DVD Escola, forneceu recentemente 50 mil aparelhos de DVD para escolas municipais e estaduais de todo o Brasil, beneficiando 853 mil professores e 21 milhões de alunos. Tais instituições necessitam de conteúdo educativo e dispõem de poucos recursos para sua aquisição. O Curta Na Escola visa atender esta necessidade através do fornecimento gratuito de conteúdo audiovisual educativo e de material didático.

Como resultado ao anúncio das tiragens gratuitas do DVDs Coleção Curta Na Escola, 19 mil escolas se cadastraram no projeto manifestando interesse em utilizar curtas-metragens no planejamento pedagógico da instituição. As escolas cadastradas junto ao Curta Na Escola reúnem mais de 10,5 milhões alunos matriculados nos níveis Fundamental

e superiores - 22% de todo o alunato nacional, pelo censo MEC/INEP.

Dentre as cadastradas, 2.800 escolas brasileiras da rede pública de ensino foram beneficiadas com os DVDs da Coleção. A seleção de instituições seguiu os seguintes critérios:

- Escolas da rede pública de ensino;
- Distribuição por todos os estados brasileiros, de acordo com a proporcionalidade de alunos matriculados (Censo MEC/Inep);
- As unidades escolares contempladas têm que possuir aparelho de DVD e computador com acesso à Internet;
- Foram priorizadas as escolas com maior número de alunos e séries.

Em contrapartida, as escolas beneficiadas se comprometeram a alimentar o banco de trocas do serviço, inserindo relatos das experiências em salas de aula e trocando experiências pedagógicas com educadores do país inteiro.

O projeto Curta na Escola é aberto à participação de professores de todo o Brasil, mesmo os de escolas que não possuam o DVD, pois a exibição dos filmes, os materiais de apoio e todas as ferramentas interativas estão disponíveis através da internet.

INSTITUTO TAMANDUÁ SYNAPSE CULTURAL

O problema brasileiro da catalogação e difusão cultural nas demais áreas - fonográfica, editorial, audiovisual, artes plásticas e artes cênicas, ainda é enorme. É difícil saber, por exemplo:

... em quais filmes atuou José Lewgoy?

... de quais livros foram adaptados os roteiros?

... quem são os autores das músicas dos filmes, em que CDs elas ocorrem?

... que outros atores participaram dos filmes, quais peças teatrais encenaram?

... como é a obra completa do artista plástico que fez a cenografia ou o cartaz de um dos filmes?

Informações interdisciplinares como estas já poderiam estar acessíveis ao público via web, em 10 minutos, e não demandar meses de trabalho de um pesquisador especialista. Porém, persiste ainda o crônico dilema da cultura brasileira: "As informações sobre a produção cultural brasileira se perdem porque ninguém as usa ou ninguém as usa porque elas se perdem?"

A experiência do Porta Curtas sugere uma resposta a esse dilema. É possível, sim, dar um salto extrapolando a biblioteconomia tradicional, de forma rica, interativa e prazerosa, democratizando de fato a circulação dos bens culturais.

O sucesso do Porta Curtas motivou a fundação, em 2005, do Instituto Tamanduá Synapse Cultural, Oscip dedicada a apoiar e promover iniciativas de catalogação e difusão, via internet, de informações rich media sobre a produção cultural brasileira.

O Instituto Tamanduá Synapse Cultural parte do princípio de que é necessário elevar o Brasil à posição de referência mundial em catalogação, pesquisa e difusão das informações sobre a produção cultural; e propõe a criação de um marco na História da produção cultural brasileira deixando para trás o Brasil "esquecido", em que predomina a sensação de estarmos sempre recomeçando, e chegar ao Brasil "lembrado", que entende a memória como um ato de soberania, afirmação, respeito, estratégia e inteligência.

Os criadores do Porta Curtas entendem que muito há por avançar, mas sabem que semearam um serviço inovador para inclusão digital e difusão cultural, que pode ser a base de construção de um projeto interdisciplinar, mais amplo, que contemple não apenas o setor de curtas-metragens, mais os demais setores – fonográfica, editorial, audiovisual, artes plásticas, artes cênicas, e assim por diante.

Repositórios institucionais:
a experiência da Universidade de Brasília

Marília Augusta de Freitas
Patrícia Nunes da Silva
Jacqueline de Farias Guimarães

INTRODUÇÃO

Com o advento da tecnologia, novos meios de comunicação surgiram, e com eles novas formas de armazenamento da informação. Nesse novo panorama, destacam-se os repositórios institucionais de acesso aberto. À medida que é facilitado o acesso à informação por meio desses repositórios, qualquer indivíduo passa a ter o direito de acessar as informações neles disponíveis. Isso porque repositórios de acesso aberto inserem-se no movimento de acesso aberto (*open access*), que permite ampliar a comunidade de usuários da informação de modo amplo e irrestrito.

Em concordância com o avanço das tecnologias e o advento das iniciativas de acesso aberto, a Biblioteca Central da Universidade de Brasília desenvolveu o Repositório Institucional da Universidade de Brasília (RIUnB) que tem como missão a gestão e disseminação da produção científica e acadêmica da Universidade. O presente trabalho mostra como foi desenvolvido o projeto do RIUnB desde sua idealização até situação atual.

REPOSITÓRIOS INSTITUCIONAIS

O desenvolvimento de Repositórios Institucionais tem-se dado amplamente, no contexto de universidades e instituições governamentais, embora mais recentemente e em escala significativamente menor.

Para Lynch (2003) no que concerne às universidades, os RIs representam uma nova estratégia que permite influenciar de maneira séria e sistemática as mudanças aceleradas que vêm ocorrendo na produção do saber e na comunicação científica.

Os RIs são, em sua maioria, usados por universidades com a preocupação de disponibilizar os resultados de pesquisas feitas por seus docentes e discentes, o que dá uma grande projeção acadêmica e permite o acesso de maneira irrestrita, rápida e gratuita por qualquer pessoa com a acesso à internet. Para Costa e Leite (2006), os repositórios têm o potencial de servir como indicadores tangíveis da qualidade de uma universidade e de demonstrar a relevância científica, social e econômica de suas atividades de pesquisa, aumentando a visibilidade, o *status* e o valor público da instituição.

A Universidade de Brasília começou a se inserir no contexto de Repositórios Institucionais em 2003, quando a professora Sely Costa, do Departamento de Ciência da Informação e Documentação (CID) participou da 7ª Conferência Internacional de Publicações Eletrônicas (Elpub), em Portugal. Naquela oportunidade, duas bibliotecárias do Massachusetts Institute of Technology (MIT) apresentaram o trabalho recém concluído sobre o DSpace, motivando a professora da UnB a desenvolver no Brasil o mesmo trabalho. Participou, igualmente, da conferência o coordenador de projetos do Instituto Brasileiro de Informação em Ciência e Tecnologia (IBICT), Dr. Hélio Kuramoto, que iniciou, também gestões no IBICT para implementação da iniciativa no Brasil.

Depois de estruturar o projeto de Repositório para a UnB em 2004, a equipe da UnB o apresentou para o Reitor, que aprovou de imediato a iniciativa. Foi firmada, então, uma parceria entre o CID, o IBICT e a Biblioteca Central da Universidade de Brasília (BCE) para desenvolvimento dos primeiros estudos sobre o DSpace e o repositório da UnB.

O DSpace é um *software* que foi desenvolvido para possibilitar a criação de Repositórios Digitais. Possui funções de captura, distribuição e preservação da produção intelectual, permitindo o gerenciamento da produção científica em qualquer tipo de material digital.

Com os estudos avançados, em 2005 o projeto foi definitivamente

implementado no servidor da BCE, com o aval da Reitoria e do Decanato de Pesquisa e Pós-Graduação (DPP) da UnB e em 2006 foram realizados treinamentos na Universidade do Minho, em Portugal, que detinha um grande conhecimento devido ao projeto que havia sido realizado com o DSpace.

REPOSITÓRIO INSTITUCIONAL DA UNIVERSIDADE DE BRASÍLIA

O Repositório Institucional da Universidade de Brasília (RIUnB) é um conjunto de serviços oferecidos pela Biblioteca Central para a gestão e disseminação da produção científica e acadêmica da Universidade de Brasília. Todo o seu conteúdo está disponível publicamente, e por estar amplamente acessível proporciona maior visibilidade e impacto da produção científica da instituição. O RIUnB tem como objetivos armazenar, preservar, divulgar e garantir acesso à produção científica e acadêmica da Universidade de Brasília em formato digital; proporcionar visibilidade à produção científica da instituição; apoiar as atividades de pesquisa e criação do conhecimento científico e apoiar o processo de ensino-aprendizagem por meio do acesso facilitado ao conhecimento.

Partindo do conceito de que Repositórios Institucionais promovem a gestão da informação científica, o RIUnB adotou, para a divulgação de um documento, que este possua as características específicas definidas por Crow e outros autores, tais como: serem documentos científicos ou academicamente orientados; produzidos, submetidos ou patrocinados pela Universidade de Brasília ou por membros da comunidade acadêmica; estarem em formato digital e prontos para serem disseminados amplamente na rede ou, conforme exceções, acessível a níveis determinados.

No RIUnB, seguindo o padrão de documentos científicos ou academicamente orientados, são permitidas submissões de tipos de documentos específicos: artigos publicados em periódicos; trabalhos publicados em anais de eventos; teses e dissertações defendidas na UnB; teses e dissertações de docentes da UnB ainda que não defendidas na própria Universidade e livros e capítulos de livros.

Organização do RIUnB

O DSpace permite que os repositórios modelem suas estruturas conforme a necessidade de cada instituição. A Universidade de Brasília adotou o sistema de comunidades, subcomunidades e coleções, conforme sua estrutura interna, ou seja, dividido por Unidades Acadêmicas. As figuras 1 e 2 exemplificam como está estruturado o RIUnB.

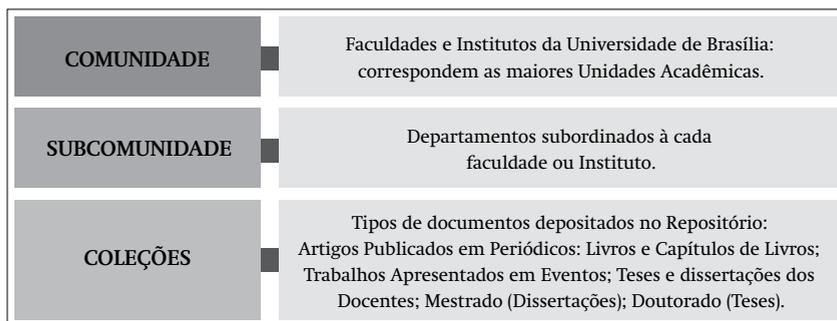



Figura 1 - Organização RIUnB

Figura 2 - Tela do Repositório Institucional da UnB

Metadados

O Repositório utiliza os metadados Qualified Dublin Core, assegurando, assim, a sua interoperabilidade, por meio do protocolo Open Archives Initiative - Protocol for Metadata Harvesting (OAI-PMH). Para garantir uma identificação persistente dos documentos arquivados no Repositório é utilizada uma URL persistente (handle)[1], isto é, uma URL que não deve mudar mesmo que o arquivo mude de servidor.

Documentação

O RIUnB possui como documentação básica uma política de funcionamento, incluindo a política de acesso aberto na qual está definido que a maioria de seus documentos está disponível para livre acesso. Inclui, também, a política de depósito obrigatório e sugestões de políticas de incentivo ao depósito obrigatório (todas em discussão, no momento, nas Câmaras de Ensino e de Pesquisa da Universidade). É importante observar, no entanto, que alguns documentos podem ter seu acesso restrito em respeito aos direitos autorais.

Além dessas políticas, foi elaborado também, um manual de uso e catalogação que prevê uma padronização na inserção dos dados no Repositório e mostra o passo a passo de uma submissão. Elaborou-se, ainda, um documento com as Perguntas Frequentes e uma ajuda, disponíveis no *website* do Repositório, a fim de esclarecer possíveis dúvidas dos usuários.

Toda essa documentação é de extrema necessidade, pois padroniza e estabelece claramente todos os objetivos e metas do serviço, evitando, assim, problemas futuros quanto à organização dos dados e documentos inseridos no Repositório. Está disponível para distribuição entre as universidades brasileiras atendidas pelo projeto do IBICT que distribuiu servidores para criação de repositórios.

---

1 O Handle System é administrado pela Corporation for National Research Initiatives (CNRI). Consulte o sítio http://www.handle.net/

Atividades Desenvolvidas

As atividades desenvolvidas pelo RIUnB estão organizadas em etapas, e, atualmente, é a Biblioteca Central que coordena e realiza todas as atividades relacionadas ao Repositório. A Biblioteca adota o modelo que consiste de duas ações básicas. A primeira, de sensibilização (visita aos departamentos para apresentar e discutir o projeto, e convidar para participar). A segunda, de ação propriamente dita (consulta aos Currículo Lattes dos professores; coleta de material na Internet, no Scielo, em páginas de revistas científicas, nos catálogos da BCE e em outros sítios, assim como junto aos próprios professores; obtenção de autorização para depósito; submissão assistida, realizada pelos bolsistas, estagiários e bibliotecários; revisão de metadados). A figura 3 mostra as etapas seguidas para o desenvolvimento das pela Biblioteca Central.

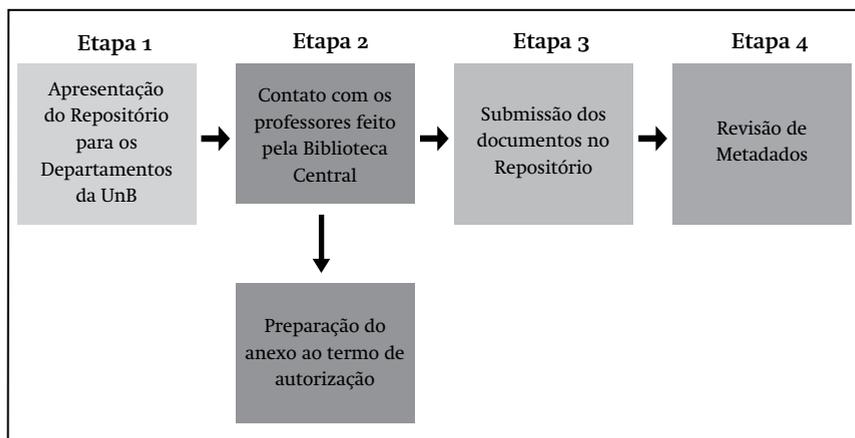



Figura 3 - Etapas de Trabalho da Biblioteca Central

A etapa 1, sensibilização e apresentação do Repositório, consiste na visita aos departamentos realizada pela diretora da Biblioteca Central, para apresentar o repositório; a visita é marcada para o dia da reunião de colegiado para que a apresentação seja feita a todos os professores.

Na etapa 2 é feita a coleta e autorização dos documentos que farão parte do RIUnB. A Biblioteca entra em contato com os professores

solicitando um agendamento de uma visita. Após agendada a visita, é preparado uma pasta contendo folder, texto explicativo, cartilha explicativa sobre o *creative commons*, termo de autorização e anexo ao termo. O anexo ao termo é preparado seguindo um modelo de planilha na qual são colocados todos os dados referentes ao documento, inclusive sua localização na *web* ou no acervo da Biblioteca. No anexo ao termo é colocado, também, a licença adotada pelo periódico. Para isso, é feita uma busca nos *websites* de periódicos eletrônicos a fim de localizar qual licença ou diretriz eles adotam para publicação em Repositórios Institucionais.

A etapa 3, submissão de documentos, é feita com base no termo de autorização, pois a Biblioteca licencia de acordo com as condições determinadas pelo autor. Todas as submissões são realizadas pela equipe da BCE.

A revisão de metadados, etapa 4, é feita exclusivamente por bibliotecários, de forma minuciosa para que não haja erro. As palavras-chave são checadas cuidadosamente, seguindo um padrão já adotado a fim de evitar duplicidade. É conferido o arquivo do documento para verificar se foi publicado de forma correta.

Estatísticas

O RIUnB tem tido altos índices de consultas e *downloads*. É possível obter esses dados através de um módulo de estatística que foi customizado pela Universidade do Minho e adaptado pela Universidade de Brasília. A figura 4 mostra os dados obtidos em 2009 e até janeiro de 2010.

O módulo de estatística é a maneira mais eficaz de apresentar o Repositório às comunidades acadêmicas, pois mostra o quanto um documento pode ser acessado e dos mais diversos países.

Na estatística é possível contemplar duas vertentes: a de *downloads* e a de consultas. Dentro de cada uma pode-se visualizar em quais países um determinado documento foi consultado e em quais foi feito o download. A figura 5 mostra a tabela de países que fizeram *download* e que consultaram determinado documento em 2009.

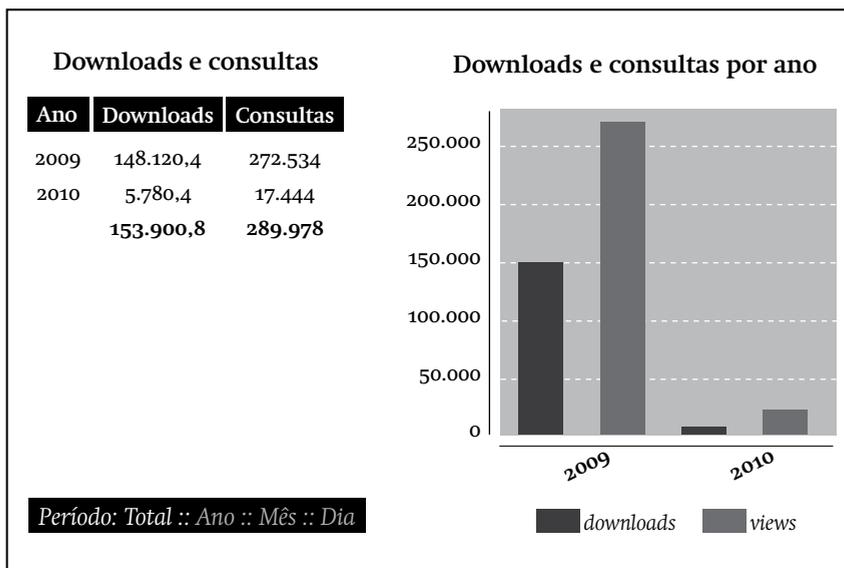



Figura 4 - Download e Consultas no Repositório em 2009 e 2010

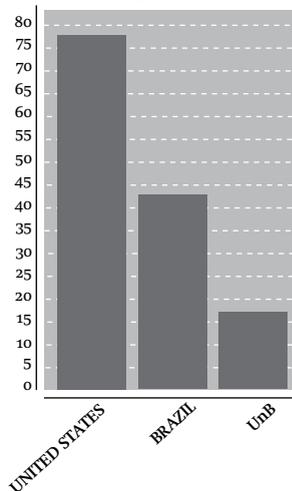



Figura 5 - Estatística de um determinado documento em 2009 e 2010

CONCLUSÕES

Os repositórios institucionais são, de fato, eficientes ferramentas para a gestão da informação e a promoção da comunicação científica principalmente em ambientes acadêmicos. A intenção deste artigo foi mostrar como a Universidade de Brasília (UnB) procurou inserir-se nesse novo contexto de informação digital e tecnologias que permitem a divulgação do conhecimento científico. O Repositório Institucional da Universidade de Brasília (RIUnB) foi criado, de fato, para atender uma demanda latente que buscava uma forma acessível de obter a produção científica da UnB.

Com o lançamento do RIUnB, os trabalhos foram intensificados, focando principalmente o contato com os docentes da Universidade. A receptividade tem sido positiva e os docentes têm mostrado bastante interesse em participar do projeto. Entretanto, ainda há, como no resto do mundo, resistências quanto ao processo de auto-arquivamento, pois os docentes alegam falta de tempo e recursos humanos para realizar essa atividade. Devido a isso, a Biblioteca Central (BCE) assumiu todas as atividades que envolvem o RIUnB, desde o contato com os docentes, passando pelas submissões até a revisão de metadados.

A próxima etapa do trabalho da BCE é intensificar a divulgação do RIUnB traçando um plano de marketing. A intenção da BCE é contactar o maior número possível de docentes, conscientizando-os da importância e necessidade de manter o RIUnB atualizado com suas produções científicas.

REFERÊNCIAS

COSTA, Sely M. de S. ; LEITE, Fernando César Lima. Repositórios institucionais: potencial para maximizar o acesso e o impacto da pesquisa em universidades. In: CONFERÊNCIA IBEROAMERICANA DE PUBLICAÇÕES ELETRÔNICAS NO CONTEXTO DA COMUNICAÇÃO CIENTÍFICA, 1., 2006, Brasília. **Anais...** Brasília: Universidade de Brasília, 2006.

GASQUE, Kelley Cristine Gonçalves Dias. **O pensamento reflexivo na busca e no uso da informação na comunicação científica.** 2008. Tese (Doutorado

em Ciência da Informação) – Universidade de Brasília, Brasília .

KURAMOTO, Hélio. **IBICT estimula adoção dos Open Archives no Brasil.** Instituto Brasileiro de Informação em Ciência e Tecnologia. 2005. Disponível em: http://www.ibict.br/. Acesso em: 02 jan. 2010.

LEITE, Fernando César Lima. **Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto.** Brasília: IBICT, 2009. v. 1. 124 p.

LEITE, Fernando César Lima; COSTA, Sely M. de S. Gestão do conhecimento científico: proposta de um modelo conceitual com base em processos de comunicação científica. **Ci. Inf.**, Brasília, v. 36, n. 1, p. 92-107, jan./abr. 2007. Disponível em: < http://revista.ibict.br/ciinf/index.php/ciinf/article/view/838/689>. Acesso em: 11 jan. 2010.

LYNCH, Clifford. Institutional repositories: essential infrastructure for scholarship in the digital age.**ARL Bimonthly Report**, n.226, fev. 2003

MORENO, Fernanda Passini; LEITE, Fernando César Lima.; ARELLANO, Miguel Ángel Márdero. Acesso livre a publicações e repositórios digitais em ciência da informação no Brasil. **Perspect. ciênc. inf.**, v.11, n.1, p. 82-94, jan./abr. 2006. Disponível em: < http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-99362006000100007>. Acesso em: 11 jan. 2010.

SUBER, Peter. **A very brief introduction to Open Access.** Disponível em: http://www.earlham.edu/~peters/fos/brief.htm. Acesso em: 11 jan. 2010

TRIGUEIRO, Osvaldo. O estudo científico da comunicação: avanços teóricos e metodológicos ensejados pela Escola Latino-americana. **PCLA**, v.2, n.2, jan./mar.2001. Disponível em: < http://www2.metodista.br/unesco/PCLA/revista6/artigo%206-3.htm>. Acesso em: 11 jan. 2010.

Memória da pesquisa, criação e inovação nas IFES:
ri para artes híbridas

Lidia Brandão Toutain[*]
Rosa Inês de Novais Cordeiro[**]

INTRODUÇÃO

A Universidade Federal da Bahia para funcionamento das suas atividades acadêmicas e administrativas designa no seu Regimento Geral como atividades essenciais o ensino, a pesquisa, a criação e a inovação, a extensão universitária e a administração institucional e acadêmica.

São consideradas atividades de pesquisa, criação e inovação aquelas que compreendem a concepção, participação, realização e coordenação dos projetos e programas geradores de conhecimento científico e tecnológico e de criação artística e cultural, nas seguintes modalidades:

> I – estudos filosóficos, teóricos, históricos ou políticos;
> II – pesquisas de campo, estudos etnográficos e similares;
> III – pesquisa-ação, intervenções comunitárias e similares;
> IV – operação de laboratórios e observatórios;
> V – desenvolvimento metodológico e instrumental de pesquisa;
> VI – pesquisa operacional e de processos institucionais;
> VII – estudos de processos de criação científica, tecnológica, artística e cultural;
> VIII – concepção e elaboração de obras de arte e similares;
> IX – investigação e experimentação em ciências básicas;

* Profª e diretora do ICI/ UFBA, Drª Filosofia (Universidad de León – Espanha)

** Profª do Dept. Ciência da Informação da UFF, Drª Comunicação e Cultura (UFRJ)

> X – outras atividades de pesquisa, criação e inovação, definidas em norma específica pelo
> Conselho Acadêmico competente. (UNIVERSIDADE FEDERAL DA BAHIA, 2010)

Baseado nestes princípios do Regimento da UFBA foi definida a política de RI (portaria n. 149 /2009) com a visão de integrar as unidades que a constitui na busca de atingir o desenvolvimento e o crescimento, respeitando as características individuais das unidades, buscando mecanismo para ampliar a visibilidade da instituição e somar esforços dentro e fora da instituição. Esta política é mais visível na área cultural, ou seja, através da música, dos livros, das artes visuais, da dança, do teatro e do cinema, onde é possível recriar linguagens, romper barreiras e superar distância e permite integração, convergência cultural e um sentido de identidade, que representa o conjunto de ações e medidas de desenvolvimento da ciência e atividade artística, visando a contribuição da UFBA para a investigação mundial e inovação entre os setores produtivos, centros de pesquisa e universidades. Estes são os princípios que inspiram a portaria da UFBA como garantia de uma política científica e das artes integral, coerente e rigorosa em seus distintos níveis de planejamento, execução e seguimento com a finalidade de obter, o necessário incremento de recursos para a investigação, a rentabilidade científica e cultural, social e econômica adequadas as suas necessidades, a fim de manter-se na vanguarda tecnológica com a especificação de política interna e externa para a criação do repositório digital.

Como instrumento de caráter institucional a portaria cria uma comissão interdisciplinar que deverá programar as atividades investigadoras, pesquisa das unidades que formam a UFBA tomando como base o plano nacional do Ministério de Ciência e Tecnologia/IBICT com a aplicação do sistema DSpace que possibilita o gerenciamento de diferentes recursos no formato digital, para uma infraestrutura de acesso livre e criação de metadados.

Segundo Viana, nas palavras de Blattmann e Weber (2008, p. 471),

> [...] a estrutura do DSpace provê um modelo de informação organizacional baseado em ´comunidades` e coleções, o qual pode ser configurado de modo a refletir todo o conjunto de unidades administrativas de uma instituição. [...] Suporta os mais variados tipos de formatos de arquivos digitais, incluindo textos, som e imagem.

Busca-se disseminar a produção cientifica, cultural e artística professores\pesquisadores universitários que realizam a investigação científica, no seu campo de especialização, contribuindo para a visibilidade institucional, a partir do acesso livre aos saberes nas ciências e humanidades que possibilita recuperar, armazenar, disponibilizar e divulgar.

De acordo com Alvarenga (2008, p. 81) os avanços tecnológicos construídos a partir do uso de instrumentos eletrônicos concebidos com o objetivo de registrar e comunicar pensamentos, ideias, imagens e sons, possibilitou que conteúdos fossem disponbilizados a um segmento determinado, porém ilimitado, de pessoas dispersas onde quer que a internet e a plataforma WWW possam alcançar.

Na declaração de Berlim (2003) o acesso livre ao conhecimento é definido como

> [...]um procedimento vantajoso que requer o empenho ativo do todo e qualquer indivíduo que produza conhecimento científico, ou seja, detentor de patrimônio cultural. Contribuições em acesso livre incluem resultados de investigação científica, originais, dados não processados e metadados, fontes originais, representações digitais de materiais pictóricos, gráficos e materiais acadêmicos multimídia.

Desse modo, os conteúdos para serem disponibilizados devem: ter concordância do autor (es) ou organizador (es) e o depósito da versão completa da obra, num formato eletrônico.

CONSTRUÇÃO E PLANEJAMENTO DO RI APLICADO AS ARTES HÍBRIDAS

Preliminarmente é interessante apresentar o conceito de "arte híbrida" aqui adotado como:

> [...] um tipo de arte resultante da constante pesquisa e do experimentalismo da pluralidade de expressões, é um produto do desenvolvimento desses ´fenômenos`. Consiste na produção de arte concretizada por meio das mais variadas técnicas, materiais e suportes, ligando linguagens artísticas e caracterizada pelo não pertencimento a uma única vertente ou categoria. (SIMÃO, 2008, p. 9) .

Continuando, afirma Simão, (2009) que nas artes visuais já não há limite entre pinturas, desenhos, fotografias, esculturas, já que envolve outras linguagens, como o cinema, teatro, dança, o ciberespaço, a música e o design. Este conceito está em consonância com a definição de Repositório Institucional, segundo o IBICT,

> [...] são sistemas de informação que armazenam, preservam, divulgam e dão acesso à produção intelectual de comunidade universitárias. Ao fazê-lo, intervêm em duas questões estratégicas: contribuem para o aumento da visibilidade e o valor público das instituições, servindo como indicador tangível da sua qualidade; permitem a reforma do sistema de comunicação científica, expandindo o acesso aos resultados da investigação e reassumindo o controle acadêmico sobre a produção científica. (DSPACE, 2005)

Ainda, sobre a natureza dos RI, Brown (2010, p. 116-117) ressalta sobre a necessidade de um repositório institucional ter uma fração significativa do seu conteúdo gerada nos *campi* da instituição e disponibilizada para o acesso público. Acreditamos que este aspecto sobre a representatividade institucional do conteúdo deve ser visto com

cuidado nos acordos de colaboração e cooperação entre instituições. Conforme este autor, um repositório institucional conterá uma coleção de documentos de trabalho (pré ou pós-prints), poderá também incluir relatórios técnicos, manuscritos, vídeo clipes, imagens e outros, além de agregar documentos e informações de apoio administrativo para a instituição, tais como a documentação arquivística.

Diante da complexidade dos aspectos a serem levados em consideração no planejamento, buscar-se-á fazer um diagnóstico situacional das artes híbridas na UFBA, especificando os objetivos, os conteúdos existentes, os usuários dos serviços, os serviços prioritários, os recursos utilizados para a comunidade externa e interna da Universidade. É importante considerar para o planejamento as diferenças de conteúdo, de necessidades da comunidade, das tecnologias adotadas no processo de criação e a forma de comunicação do conhecimento.

Para o desenvolvimento deste diagnóstico é necessário:

a) Mapear as obras artístico/cultural das unidades, por área de conhecimento;
b) Formar equipe interdisciplinar com os produtores e pesquisadores de arte/ cultura, historiadores, museólogos, bibliotecários, arquivistas analistas de sistemas, que trabalham nas unidades de dança, teatro, belas artes, enfim profissionais das linguagens visuais, para discussão a respeito da política de RI da IFES;
c) Levantar as formas de uso e comunicação utilizadas pelas diferentes unidades da instituição a fim de sensibilizá-las para uma aplicação na gestão da informação digital *on-line.*

Esta fase preliminar busca garantir o ambiente necessário para as definições posteriores do planejamento.

No caso da equipe interdisciplinar alguns conhecimentos e habilidades são elencados por (LEITE, 2009): cada profissional envolvido deve estar apto a lidar com os processos voltados para a implementação e funcionamento do RI. Por exemplo, os analistas de sistemas devem dominar os requisitos tecnológicos necessários para instalação, configuração e customização das ferramentas. Os bibliotecários de-

vem dominar processos de gestão e ter habilidades no domínio dos métodos de identificação avaliação de necessidades de informação da comunidade, assim como técnicas e instrumentos de organização em ambiente eletrônico e familiaridade com as Tecnologias de Informação e Comunicação (TICs). Os produtores de informação, os gestores e investigadores devem estar envolvidos em todo processo e serem treinados para desenvolver políticas, práticas e procedimentos para assegurar que o repositório se torne presente nas unidades e consequentemente na instituição com o objetivo de compartilhar informação, estabelecer contatos com os pesquisadores, grupo de pesquisa, para os principais atores envolvidos na criação, registro e disseminação da informação e do conhecimento das artes híbridas, para identificar suas necessidades e uso no repositório e estimular o envolvimento com os serviços.

Outra fase do planejamento é a observação das forças externas e internas à instituição que influencia a produção de cada unidade que compõe a UFBA.

INFORMAÇÃO VISUAL, DINÂMICA E ESTÁTICA NO REPOSITÓRIO INSTITUCIONAL

Na literatura de Ciência da Informação identificam-se diferenças de uso e necessidades de informação nos vários segmentos. Na área das artes híbridas verificam-se necessidades especiais o que demanda auxílio no tratamento da informação que seja mais adequada ao contexto digital.

Blatmann e Weber (2008) orientam para que o conteúdo do repositório seja organizado em comunidades, entidades administrativas conhecidas como coleções. Explicam que cada subcomunidade pode ser uma coordenação de áreas temáticas, ou um grupo de ação e manifestação cultural que atue no repositório. Assim, os documentos serão denominados de itens e serão agrupados em coleções. Os itens seguem os padrões de metadados, do próprio *software* pautado no Dublin Core.

No âmbito das universidades são produzidos documentos únicos, de importância singular que são resultantes das práticas acadêmicas

realizadas nas graduações e pós-graduações dos diversos cursos. Entretanto, a representação e a recuperação desses documentos (itens) carecem de análise quanto a sua estrutura e conteúdo. Para acentuarmos e demonstrarmos tal afirmativa, bem como enfatizarmos a importância desses documentos que são gerados e poderão integrar os repositórios institucionais, enfocaremos uma possibilidade real de alguns cursos de cinema e audiovisual. Para isto nos inspiramos nos procedimentos que são realizados na Graduação de Cinema e Audiovisual da Universidade Federal Fluminense (UFF), quando os alunos fazem a produção de um filme de curta-metragem, constituindo, portanto, a produção filmográfica do corpo discente do Departamento de Cinema e Vídeo que é criada no ambiente institucional citado e que reforça a ideia da multiplicidade da natureza dos documentos que deverão integrar um repositório institucional das IFES e, aqui, no âmbito artístico, cultural e social.

Então, nossa intenção é abordar a produção de filmes de curta-metragem e a sua família de documentos como futuros objetos digitais multimidiáticos integrantes de repositórios institucionais.

Para a realização de um filme diversas ações são pensadas e desenvolvidas, e em consequência, documentos que registrarão a tomada de decisões que serão implementadas nas etapas de pré-produção/produção/pós-produção do produto audiovisual são gerados (digitais ou não digitais).

Pode-se distinguir as seguintes atividades em cada etapa:

- Pré-produção: roteiro; despesas da pré-produção (transporte, gratificações, testes, outros); escolha de equipes; material sensível; equipamento de filmagem; material de iluminação;
- Produção: equipes; elenco; despesas de produção (transporte, gratificação, revelação de still, outros); laboratório (revelação de negativo, copião, outros);
- Pós-produção: montagem; equipe técnica; estúdio de som; lab Na pré-produção e na produção são feitas as previsões de despesas.

De forma mais complexa, Rabley (1990) sistematiza o processo de produção propriamente dito de um filme nas seguintes etapas e

atividades: ideia, texto, recursos financeiros, pré-produção, produção, pós-produção, lançamento/marketing, comercialização. Este processo é sistematizado de forma resumida nas seguintes **etapas e ações/ atividades para a realização de um filme**:

- IDEIA: observação e/ou pesquisa as fontes informacionais/ adaptações;
- TEXTO: redação do argumento/sinopse/roteiro;
- RECURSOS FINANCEIROS: levantamento dos recursos para financiamento do projeto;
- PRÉ-PRODUÇÃO: formação do elenco e equipe técnica, locações, construção de sets, elaboração de figurinos;
- PRODUÇÃO: filmagem (não menos que 3 meses);
- PÓS-PRODUÇÃO: término da filmagem (edição; adicionar a trilha sonora ao filme; dublagem/legenda);
- LANÇAMENTO/MARKETING: permissão da exibição por certificado, promoção e lançamento, campanha publicitária para divulgação do produto;
- COMERCIALIZAÇÃO: venda do produto através de contratos.

Embora haja autores que colocam também em pré-produção a ideia, o texto e os recursos financeiros - concordamos com o esquema de Rabley e podemos denominar a fase anterior à pré-produção como "viabilização".

Nas etapas descritas acima, diversos documentos são gerados e contendo informações de grande relevância para preservação da produção filmográfica institucional, também no âmbito da filmografia regional e nacional, e o seu acesso necessita ser disponibilizado, pois essa documentação normalmente não faz parte dos acervos arquivísticos e das cinematecas brasileiras.

A seguir, demonstramos (Quadro 1) a relação entre documentos que são criados para um filme e as atividades desenvolvidas para sua realização:

DOCUMENTOS GERADOS	ETAPAS AÇÕES/ATIVIDADES
ARGUMENTO	viabilização/produção
SINOPSE	viabilização/produção
ROTEIRO	viabilização/produção
STORYBOARD	viabilização/produção
ANÁLISE TÉCNICA DO ROTEIRO	produção
CRONOGRAMA OU PLANO DE DESEMBOLSO	produção
PLANO DE PREPARAÇÃO	produção
PLANO DE FILMAGEM	produção
PLANO DE FINALIZAÇÃO	produção e pós-produção
FOLHA DE CONTINUIDADE	Produção
ORDEM DO DIA	Produção
BOLETIM DE CÂMERA	produção e pós-produção
BOLETIM DE SOM	produção e pós-produção
CONTROLE DE ESTOQUE DE MATERIAL SENSÍVEL DE SOM E IMAGEM	Produção
RELAÇÃO DE EQUIPAMENTO DE SOM, IMAGEM E MAQUINARIA ELÉTRICA	Produção
CONTROLE DE MATERIAL DE CENOGRAFIA, FIGURINO, MAQUIAGEM E EFEITO	Produção
CONTROLE DE LABORATÓRIO	produção e pós-produção
TEXTO DE DUBLAGEM (não é feito quando o filme é rodado em som direto)	pós-produção
PLANO DE PRODUÇÃO DE DUBLAGEM (não é feito quando o filme é rodado em som direto)	pós-produção
AGENDA DE DUBLAGEM (feito dentro de plano de produção de dublagem)	pós-produção
CONTROLE DE LOOPS PARA DUBLAGEM	pós-produção
MAPA DE MIXAGEM	pós-produção
MAPA DE TRUCAGEM	pós-produção
CERTIFICADO DE CENSURA	pós-produção
PRESS RELEASE ((RELEASE)	pós-produção/lançamento
TRAILER	pós-produção/lançamento
CARTAZ	pós-produção/lançamento
CARTAZETE	pós-produção/lançamento
FOTO-CARTAZ	pós-produção/lançamento
FOTO DE CENA (FOTOGRAFIA DE CENA/STILL)	produção/pós-produção/ lançamento
OUT-DOOR	pós-produção/lançamento
DISPLAY	pós-produção/lançamento
CONTRATOS DE COMERCIALIZAÇÃO	pós-produção
BORDERÔ SOBRE A RENDA DOS FILMES	pós-produção

Quadro 1 – Documento gerado e etapa/atividade de realização do filme

O potencial informativo desses documentos deve ser analisado como fonte de pesquisa, tendo em vista a descrição e a recuperação de algumas categorias (atores, personagens principais etc) como possíveis pontos de acesso para a busca do conteúdo do documento. A seguir, demonstramos tal proposta.

PONTOS DE ACESSO (informações "de":)	**DOCUMENTO GERADO PARA O FILME (responde a demanda solicitada pelos pontos de acesso)**
ATORES	Análise técnica
	Plano de filmagem
	Ordem do dia
PERSONAGENS PRINCIPAIS	Argumento
	Sinopse
	Roteiro
	Análise Técnica
	Plano de Filmagem
	Ordem do Dia
FIGURANTES	Roteiro
	Análise técnica
	Plano de filmagem
	Ordem do dia
CONFLITO-MATRIZ	Argumento
	Sinopse
	Roteiro
RESUMO DA HISTÓRIA (resumo do filme)	Argumento
	Sinopse
HISTÓRIA NA INTEGRA	Análise técnica
	Folha de continuidade
	Roteiro
ESPAÇO/TEMPO	Argumento
	Sinopse
	Folha de continuidade
	Ordem do dia
LOCALIZAÇÃO/AMBIENTE	Roteiro
	Análise técnica
	Folha de continuidade
	Ordem do dia
CENÁRIO	Análise Técnica
	Plano de Filmagem
	Folha de Continuidade
	Ordem do Dia
OBJETOS DE CENA / MÓVEIS E ACESSÓRIOS	Análise Técnica
	Ordem do Dia
FIGURINO / MAQUIAGEM / CABELO	Análise Técnica
	Folha de Continuidade
	Ordem do Dia
TRILHA SONORA	Análise Técnica
	Plano de Filmagem
RITMO DA DANÇA (para filme musical)	Análise Técnica do Roteiro

Quadro 2 – Pontos de acesso (informações "de") e documentos

A partir do quadro acima, os documentos gerados para o filme poderão responder, por exemplo, as seguintes perguntas:

- Em quais documentos obtenho informações sobre o **resumo da história do filme**? Resposta: obtenho informações precisas na **sinopse** e no **roteiro**.
- Em quais documentos obtenho informações sobre **objetos de cena/móveis** e **acessórios** das sequências do filme? Resposta: obtenho informações precisas na **análise técnica** e na **ordem do dia** do filme.
- Em quais documentos obtenho informações sobre os cenários? Resposta: obtenho informações precisas na **análise técnica**, **plano de filmagem**, **folha de continuidade** e na **ordem do dia** do filme.

Tendo em vista o exposto, pensamos que os repositórios institucionais das IFES devem considerar a possibilidade de agregar aos seus conteúdos os filmes e a família de seus documentos como uma fonte de informação inovadora e de substancial importância para a pesquisa social, cultural e das artes híbridas.

CONSIDERAÇÕES FINAIS

A implantação do RI para "artes híbridas" favorecerá as atividade de pesquisa, criação e inovação, que compreendem a concepção, participação, realização e coordenação dos projetos e programas geradores de conhecimento científico e tecnológico e de criação artística e cultural da IFES/UFBA.

Os documentos que são gerados nas IFES para as suas atividades de ensino, pesquisa e extensão apresentam informações únicas e mostram a organicidade institucional através da articulação das atividades desenvolvidas e dos seus registros, sejam documentos digitais ou não. Acreditamos que esses registros encontram-se silenciados pelo não uso e, portanto, poderão ser incluídos nos repositórios das IFES e acessados como uma fonte de pesquisa diferenciada e resultante de determinado contexto social, científico e artístico.

Por fim, gostaríamos de finalizar este texto destacando três aspectos que são citados por Brown (2010, p. 123) e que dizem respeito às recomendações da Associação de Bibliotecas de Pesquisa – Association of Research Libraries (ARL) – para o serviço de repositório digital quando indicam:

- a importância de estratégias básicas para o desenvolvimento de serviços baseados em uma avaliação substantiva das necessidades locais ao invés da replicação cega de trabalhos feitos por outras instituições;
- o compromisso com as questões fundamentais da política local e encorajamento da participação do engajamento institucional com as questões das políticas nacional e internacional;
- a definição de um escopo de responsabilidades para guiar o desenvolvimento dos serviços de repositórios para as variadas formas de conteúdos.

REFERÊNCIAS

ALVARENGA, L. Organização da Informação nas bibliotecas digitais. In: KURAMOTO, H.; NAVES, M. *Organização da informação princípio e tendência*. Brasília: Briquet de Lemos, 2006.

BLATTMANN, U; WEBER, C.. DSpace como repositório digital na organização. *Revista ACB.* Florianópolis, v.13, n. 2, p. 467-85, jul./dez, 2008.

BROWN, D. J. Repositories and journals: are they in conflict? *Aslib Proceedings*, v. 62, n. 2, p. 112-143, jan. 2010.

CORDEIRO, R. I. de N. *Informação e movimento:* uma ciência da arte fílmica. Niterói: UFF, Programa de Pós-Graduação em Ciência da Arte, 2000.

DECLARAÇÃO de Berlin sobre acesso livre ao conhecimento nas ciências e humanidades. In: CONFERENCE ON OPEN ACCESS TO KNOWLEDGE IN THE SCIENCES AND HUMANITIES, Berlin, 2003.

Disponível em: <http://oa.mpg.de/openaccess-berlin/BerlinDeclaration_pt.pdf>.

DSPACE - REPOSITÓRIOS INSTITUCIONAIS. *Glossário*. Brasília: IBICT, 2005. Disponível em: < HTTP://dspace.ibict.br/index.php>.

LEITE, F. *Como gerenciar e ampliar a visibilidade da informação científica brasileira:* repositório institucional de espaço aberto. Brasília: IBICT,2009.

RABLEY, S. *The Cinema Dossier*. Singapore: Macmillan, 1990.

SIMÃO, S. M. *Arte híbrida:* entre o pictórico e o fotográfico. São Paulo: UNESP, 2008.

TOMAEL, M. I. (Org.). *Fontes de informação na internet*. Londrina: EDUEL, 2008.

UNIVERSIDADE FEDERAL DA BAHIA. *Regimento Geral da Universidade Federal da Bahia*. Salvador, março, 2010.

______. *Portaria n. 149/2009* política de repositório institucional. Salvador: Gabinete do Reitor, 2010.

Sobre os autores

AIDA VARELA VARELA

Doutora em Ciências da Informação pela Universidade de Brasília (2003). Atualmente é professor adjunto da Universidade Federal da Bahia. Tem experiência na área de Ciência da Informação, com ênfase em Teoria da Informação e Teoria de Aprendizagem atuando principalmente nos seguintes temas: mediação e modificabilidade cognitiva estrutural. varela@ufba.br

CARLOS HENRIQUE MARCONDES

Doutor em Ciência da Informação pela Universidade Federal do Rio de Janeiro (1998). Atualmente é professor associado da Universidade Federal Fluminense. É também consultor em tecnologia da Fundação Oswaldo Cruz, pesquisador 1D do Conselho Nacional de Desenvolvimento Científico e Tecnológico e membro do CTC do Instituto Brasileiro de Informação em Ciência e Tecnologia. Membro de corpo editorial da Perspectivas em Ciência da Informação e revisor dos seguintes periódicos: Revista Eletrônica de Comunicação, Informação e Inovação em Saúde, Revista Brasileira de Pós-Graduação, Perspectivas em Ciência da Informação, Ciência da Informação (Impresso), Transinformação e PontodeAcesso. Tem experiência na área de Ciência da Informação com ênfase em Arquivologia atuando principalmente nos seguintes temas: políticas públicas de informação em C&T, planejamento de ICT, Brasil. marcon@vm.uff.br

EDSON RAMIRO LUCAS FILHO

Graduando do curso de Bacharelado em Ciência da Computação da Universidade Federal do Paraná. Faz iniciação científica no grupo de Banco de Dados do Centro de Computação Científica e Software Livre (C3SL) sobre Preservação Digital e Banco de dados NoSQL. erlf07@inf.ufpr.br

FLAVIA GOULART MOTA GARCIA ROSA

Doutoranda em Comunicação no Programa Multidisciplinar de Pós-Graduação em Cultura e Sociedade da Universidade Federal da Bahia (UFBA). É professor adjunto e diretora da editora da UFBA e presidente da Associação Brasileira de Editoras Universitárias (ABEU). Tem experiência na área de Comunicação, com ênfase em comunicação científica e editoração, atuando principalmente nos seguintes temas: editoração universitária, universidade, editoração cultural e científica e design gráfico editorial. fflaviarosa@gmail.com

IGOR BARAÚNA GUIMARÃES

Mestrando em Ciência da Informação pela Universidade Federal da Bahia (UFBA). Foi bolsista de Iniciação Científica pela FAPESB, desenvolvendo pesquisa no Instituto de Ciência da Informação da UFBA, no período 2007-2008. Atua como Reporter de Televisão no Instituto de Radiodifusão Educativa da Bahia (IRDEB), lotado na TV Educativa da Bahia. igor.guimaraes@hotmail.com

JACQUELINE DE FARIA SILVA GUIMARÃES

Especialista em Gestão do Conhecimento pela Universidade Gama Filho. Atuação: participação na implantação da Biblioteca Digital Ju-

rídica do Superior Tribunal de Justiça – BDJur/STJ (2004) e integrante da equipe BDJur/STJ (2004-2008). Bibliotecária responsável pelo Repositório Institucional da UnB. jacqueline.guimaraes@bce.unb.br

LIDIA MARIA BATISTA BRANDAO TOUTAIN

Doutora em Filosofia pela Universidad de León – España (2003), en el programa de Comunicación, Acción y Conocimiento. Membro do Conselho Universitário da Universidade Federal da Bahia (UFBA), Diretora do Instituto de Ciência da Informação e professor adjunto do Departamento de Fundamentos e Processos Informacionais atuando no Programa de Pós-Graduação em Ciência da Informação da UFBA. Tem experiência na área de Ciência da Informação, com ênfase em Teoria da Informação, atuando principalmente nos seguintes temas: epistemologia da comunicação, informação e imagem, representação e informação, semiótica da imagen, semiótica da cor. Coordena o Grupo Interinstitucional de Processos Semiótico e de Design (GIPSD) cadastrado no CNPq. lbrandao@ufba.br

LIGIA ELIANA SETENARESKI

Especialista em Administração Universitária (1997) pela Universidade Federal do Paraná. Ocupa atualmente o cargo de Diretora do Sistema de Bibliotecas da Universidade Federal do Paraná (SIBI/UFPR), atuando principalmente nos seguintes temas: bibliotecas digitais, aquisição de material bibliográfico, desenvolvimento de coleções, administração de bibliotecas e informatização de bibliotecas. ligia@ufpr.br

LUCIO FOLTRAN DA SILVA

Mestrando em Ciência da Computacao pela Universidade Federal do Paraná. Faz parte da equipe do Centro de Computação Científica e

Software Livre da UFPR, onde atua como bolsista responsável pelo Sistema de Biblioteca Digital desde 2008. luciof@c3sl.ufpr.br

LUIS FERNANDO SAYÃO

Doutor em Ciência da Informação pela Universidade Federal do Rio de Janeiro (1994). Trabalha desde 1980 na Comissão Nacional de Energia Nuclear onde já exerceu os cargos de: chefe do Centro de Informções Nucleares (CIN); chefe da Divisão de Tecnologia da Informação; coordenador-geral de Informática; representante do Brasil no INIS - International Nuclear Information System (AIEA/ONU); coordenador-geral da RRIAN - Red Regional de Información en el Área Nuclear. É membro do Câmara Técnica de Documentos Eletrônicos do CONARQ - Conselho Nacional de Arquivos e membro do Comitê Técnico-Científico do IBICT - Instituto Brasileiro de Informação em Ciência e Tecnologia e da Comissão de Ensino da CNEN. Tem como áreas de interesse: bibliotecas digitais, publicações eletrônicas, interoperabilidade, bases de dados e preservação digital. lsayao@cnen.gov.br

MARCOS SFAIR SUNYE

Doutor em Banco de dados pela Université de Bourgogne (1993). Realizou Pós Doutorado no Laboratório de Informática de Paris V (LIP6) da Université Paris VI (Jussieu) em 2006/2007. Atualmente é professor associado e membro do Conselho de Ensino e Pesquisa da Universidade Federal do Paraná (UFPR). É um dos líderes da área de bibliotecas junto do projeto CT-INFRA/Finep. É responsável pela criação da biblioteca digital da UFPR que reune 32 revistas eletrônicas, todas as teses e dissertaçoes desde 2004 e os vídeos produzidos pela UFPR. É membro da Comissão de Sistemas de Informação e Telecomunicações do Estado do Paraná. Preside o Comitê Gestor da rede COMEP, projeto da RNP de implantação de fibras óticas na regiao metropolitana de Curitiba. Coordena o projeto GHI que desenvolve um sistema de código livre

de Gestão Hospitalar. Tem experiência na área de Ciência da Computação, com ênfase em banco de dados, atuando principalmente nos seguintes temas: banco de dados, integração, análise de desempenho e bilbiotecas digitais. sunye@c3sl.ufpr.br

MARILENE LOBO ABREU BARBOSA

Possui o Diplôme dÉtudes Approfondies en Information Scientifique et Technique, pela Université Aix-Marseille III e mestrado em Ciência da Informação pelo Instituto de Ciencia da Informação da Universidade Federal da Bahia (2005), onde atualmente leciona. Tem experiência na área de Ciência da Informação, com ênfase em Biblioteconomia, atuando principalmente nos temas: gestão da informação e de serviços de informação, inteligência competitiva, competências, formação profissional e tecnologias de informação e comunicação. marilene@ufba.br

MARÍLIA AUGUSTA DE FREITAS

Mestre em Ciências da Informação pela Universidade de Brasília (2010). Atualmente é bibliotecária da Universidade de Brasília, chefe do Serviço de Gerenciamento da Informação Digital, responsável pela supervisão das atividades do Repositório Institucional, da Biblioteca de Teses e Dissertações e da Biblioteca Digital de Monografias. Coordena também o Diretório de Periódicos Acadêmicos da Universidade de Brasília. marilia@bce.unb.br

PATRÍCIA NUNES DA SILVA

Mestranda em Ciência da Informação pela Universidade de Brasília e chefe da Divisão de Periódicos da Biblioteca Central da Universidade de Brasília. Atualmente realiza pesquisas nas áreas de Arquitetura da

Informação, Organização do conhecimento e Organização social do conhecimento. patricia@bce.unb.br

ROSA INÊS DE NOVAIS CORDEIRO

É doutora em Comunicação e Cultura pela Universidade Federal do Rio de Janeiro (UFRJ/1998) e realizou seu pós-doutorado no Instituto de Psicologia da UFRJ (2003). É professora da Universidade Federal Fluminense (UFF) onde desenvolve atividades ligadas ao ensino, à pesquisa e a extensão no Instituto de Arte e Comunicação Social, no Departamento de Ciência da Informação e no Laboratório de Investigação Audiovisual. Atua, desde 1983, no ensino das graduações de Arquivologia; Biblioteconomia e Documentação da UFF. Desde 2000 atua também na docência da pós-graduação na UFF. Colabora também com o Mestrado em Ciência da Informação da UFF e com a subárea Ciência da Informação e Justiça Administrativa, no Mestrado Profissional em Justiça Administrativa da UFF. Na especialização participa do corpo docente do curso em Planejamento, Organização e Direção de Arquivos (UFF). Desde agosto de 2006 é coordenadora do Programa de Pós-Graduação (Mestrado e Doutorado) em Ciência da Informação, convênio UFF e IBICT. São seus principais domínios de atuação a análise e a representação do conteúdo dos documentos de diversas naturezas (audiovisuais e imagens em diferentes estoques de informação) e a geração e a gestão do conhecimento em Direito.

SELY MARIA DE SOUZA COSTA

Doutora em Ciência da Informação - Loughborough University (1999). Atualmente é professor adjunto da Universidade de Brasília, diretora da Biblioteca Central da Universidade de Brasília, membro do Conselho Editorial da Editora da Universidade de Brasília e professora visitante da Universidade do Minho / Departamento de Sistemas de Informação. Coordena o projeto de repositório institucional da UnB, em parceria

entre o CID e a BCE, assim como o Ibict. É, ainda, parecerista da várias revistas brasileiras. Sua experiência na área de Ciência da Informação é em Processos da Comunicação, atuando principalmente nos seguintes temas: publicações eletrônicas, comunicação científica, acesso aberto e comunicação organizacional. Realiza também estudos nas áreas de Gestão da Informação e de Gestão do Conhecimento (organizacional e científico) e trabalha com a Metodologia de Sistemas Flexíveis (Soft Systems Methodology) como método de pesquisa. selmar@unb.br

VANESSA DE A. SOUZA

Possui Pós-Graduação em Gestão de Negócios – IBMEC. Instituição: Instituto Tamanduá Synapse Cultural, Oscip. Atua como Coordenadora dos projetos Porta Curtas e Curtas Na Escola. vanessa@portacurtas.com.br

VERA LUCIA DOYLE LOUZADA DE MATTOS DODEBEI

Doutora em Comunicação e Cultura pela Universidade Federal do Rio de Janeiro (UFRJ). Professora associada da UFRJ no Programa de Pós-Graduação em Memória Social (mestrado e doutorado). Vice-Presidente da ISKO - Brasil (International Society of Knowledge Organization - Capítulo brasileiro). Coordenadora do GT10 - Informação e Memória, da Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação - ANCIB. Vice-Coordenadora do Programa de Pós-Graduação em Memória Social da UNIRIO. Desenvolve projetos nas áreas de Ciências Sociais, Humanas e Ciência da Informação, com ênfase em Cultura Digital, Memória e Patrimônio. Pesquisadora no Brazilian Institute of Web Science Research. Investiga as relações entre informação, memória e virtualidade; organização do conhecimento, linguagens documentárias, tesauros e redes conceituais virtuais; memória e imagem; patrimônio digital e memoria social. dodebei@gmail.com